

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2003, 12:51:58 ; Search time 8181 Seconds
(without alignments)
11586.339 Million cell updates/sec

Title: US-09-817-198b-1

Perfect score: 3257
Sequence: 1 tgcgcgtgcccgcgcgcag.....aaaaaaaaaaaaaaaaaa 3257

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.:

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pi.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pi.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_wam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2896.6	82.8	190517	9	CNS01DX4
2	1192.8	36.6	7924	6	AX347040 Human chr
3	1192.8	36.6	7924	6	AX347040 Sequence
4	1160.8	35.6	7924	6	AX348456 Sequence
5	1160.8	35.6	7924	6	AX347041 Sequence
6	1000.2	30.7	1054	6	AX348457 Sequence
7	866.2	26.6	895	6	AX399903 Sequence
8	813	25.0	3139	10	AX014147 Sequence
9	577.6	17.7	945	10	BC027769 Mus muscu
10	477.4	14.7	481	6	M83679 Sprague-Daw
11	438.8	13.5	463	6	AX396088 Sequence
12	423.6	13.0	162495	2	AX334820 Sequence
13	416.4	12.8	2560	10	AC096084 Rattus no
14	378.4	11.6	257226	2	BC013790 Mus muscu
15	307	9.4	313	6	AC124556 Mus muscu
16	234.2	7.2	2048	9	AX341683 Sequence
17	234.2	7.2	2818	9	EC002977 Homo sapi
18	230.6	7.1	1980	9	AK025165 Homo sapi
19	229	7.0	624	9	S53268 Homo sapien
20	229	7.0	660	9	AF498943 Homo sapi
21	227.4	7.0	638	10	X56741 H.sapiens m
22	224.2	6.9	1337	10	S53270 MEL-RAS-rel
23	222.6	6.8	760	4	BC019990 Mus muscu
24	218.2	6.7	765	5	X36385 Canine rab8
25	217.8	6.7	624	6	M38391 Discopyge o
26	217.8	6.7	1128	9	AX236078 Sequence
27	217.8	6.7	1161	6	BC020654 Homo sapi
28	217.8	6.7	1265	9	AX236076 Sequence
29	217.8	6.7	2497	6	AB038995 Homo sapi
30	217.8	6.7	2497	6	AX285074 Sequence
31	217.8	6.7	2497	6	AX285080 Sequence
32	217.8	6.7	2497	6	AX285089 Sequence
33	217.8	6.7	3077	9	AX285091 Sequence
34	216.2	6.6	3110	9	AK001111 Homo sapi
35	211.4	6.5	740	10	AL833365 Homo sapi
36	203.2	6.2	991	10	U53475 Rattus norv
37	201.8	6.2	897	10	M83677 Sprague-Daw
38	198.6	6.1	110000	2	AF035646 Mus muscu
39	198.6	6.1	110000	2	AC048347 Homo sapi
40	198.6	6.1	302718	2	Continuation (2 of
41	198.6	6.1	1029	9	AC027292 Homo sapi
42	198.4	6.1	3533	9	BC000896 Homo sapi
43	196.8	6.0	3521	9	AK023223 Homo sapi
44	196.6	6.0	2048	3	AL136650 Homo sapi
45	196.6	6.0	2210	3	AY069671 Drosophila

ALIGNMENTS

RESULT 1	CNS01DX4	190517 bp	DNA	linear	PRI 17-MAY-2001
LOCUS	Human chromosome 14				
DEFINITION	from chromosome 14 of Homo sapiens (Human), complete sequence.				
ACCESSION	AL139022				
VERSION	AL139022.4	GI:14148954			
KEYWORDS	HTG.				
SOURCE	Human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 190517)				
AUTHORS	Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,				
	Brottier,P., Cattoico,L., Barbe,V., Pelletier,E., Artiguenave,F.,				

Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mqc@nhgri.nih.gov

Akter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamín, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K.C., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Grant, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Mastello, C., Maskeri, B., Mastran, S.D., McCloskey, J. C.,
 McDowell, J., Pearson, R., Skatripp, S., Thomas, P.J., Touchman, J.W.,
 Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK File: 55 Row: d Column: 5
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

[illegible]**Location/Qualifiers**

1. .3139

/organism="Mus musculus"

```
/db_xref="taxon:10090"
```

$$/map="FVB/N"$$

```

/ map= TVD/N
/ clone= "MGC:38375 IMAGE:5345297"

```

```

/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
ductal carcinoma, 5 month old virgin mouse."

```

/clone lib="NCI CGAP Mam

```
/lab host="DH10B"
```

```
/run_host= build
/note="Vector: pCMV-SPORT6"
```

208. 846

```
200: .040
/codon start=1
```

```

/codon_start=1
/product="Unknown (protein for MGC:38375)"

```

```
/product="UNKNOWN (prote
/protein id="AAH27769.1"
```

```
/protein_id="AAH27769.1"
/db_xref="CT:20390732"

```

```
/db_xref="GI:20380/22"
/translation="MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVD
FKMKTIDVGIKVRQIWDTAGERYQTIKQYRRAQGIFLVYDSSERSYQHIMKW
VSDIDEYAPEGVOKILGNKADEOKROVGREGOOLAKKEYGMDFYETSACTNLNIKE
```

847 c 878 d 691 f

BASE C
ORIGIN

Query Match	25.0%	Score 813;	DB 10;	Length 3139;
Best Local Similarity	67.4%;	Pred. No. 2.1e-198;		
Matches 1602; Conservative	0;	Mismatches 640;	Indels 165;	Gaps 30;
QY	22	TCCCGGCGCCGCTGGCCCGCCAGTCATGCGCAAGCAGTACGATGTGCTGTTCCGGCTGCTGC	81	
DB	185	TCGGGCCCCCTTCAGTCCCTGCCATGGCGAAACAGTACGATGTGCTGTTCCGGCTACTGC	244	
QY	82	TGATCGGGGACTCCGGGGTGGGAAGACCTGCTGTCGCCGTTCCACGCAACAGAGT	141	
DB	245	TGATCGGGGACTCCGGGGTGGCAAGACATGCTGCTGTCGCCGTTCCACGCAACAGAGT	304	
QY	142	TCCACTCTCGGCACATCTCCACCATCGGTGTGACTTTAAGATGAAGACCATAGAGGTAG	201	
DB	305	TCCACTCTCGCATATCTCCACCATCGGTGTGACTTTAAGATGAAGACTATCGATGTAG	364	
QY	202	ACGGCATCAAGTGGCGGATACAGATCGGACACTGCAGGCGAGGAGAGATACCAGACCA	261	
DB	365	ACGGCATCAAGTGAGAAATACAGATTTGGGACACAGCAGGCGAGGAGGTACCAAGACTA	424	
QY	262	TCACAAGCAGTACTATCGCGGGGCCAGGGGATATTTTGGTCTATGACATATAGCAGCG	321	
DB	425	TCACAAGCAGTACTATCGCGGAGCCCGCGGATATTTTGTAGTCTACGACATTAGCASTG	484	
QY	322	AGCGCTTTTACCAGCACATCATGAAGTGGCTCAGTGACGTGGATGAGTACGCAACAGAG	381	
DB	485	AGCGCTCTTATCAGCATATCATGAAGTGGCTCAGTGACGTGGATGAGTACGCTCCAGAAG	544	
QY	382	GCCTCCAGAGATCCCTATTTGGCAATTAAGCTGATGAGGAGCGAAGACGCCAGGTGGAA	441	
DB	545	GAGTCCAGAAGATCCCTAATTGGGAATTAAGCTGATGAAGACGAAGACGCCAGGTGGGA	604	
QY	442	GAGAGCAAGGCGACAGCTCGCGCAAGGAGTATGSCATGGACTTCTATGAAACAAGTGCCT	501	


```

Db      1625 CGATGGGTGCGAGGAT--TAAGTCACCTCTGTCTCTACCTCCCATGCCCTC----- 1675
Qy      1574 CAATTCCTCCACTGGCTCCCAAAATTAAGATGTGGACCAAGGCTGTGGTACTCC 1633
Db      1676 ----TCTCCTCACCTGGATCACAATTAAGATGGGATCGAGGCTGAGGACAC-CA 1730
Qy      1634 AGGGCAGGAGAGCCCTGGGTCAGTGACACTGTAGGACCAACCCAGTCACTCCACAAG 1693
Db      1731 AAGGGCAGGGAGCCCTGGGTCAC--GATACTGTATAGTCAGCTGTGAATCCAGATAG 1788
Qy      1694 GGGAGCAATTTGAAATGAAGACTAGCTCTCTATGATGAGTTAAGAGCAGGAGAGCT 1753
Db      1789 AGGCAAGCGGATCCGAAGGCGGCTCTAGAACAGGTTAGACCAAGAGCAGAGG 1848
Qy      1754 GGCCAGGACAGCAGTTTGCACAGCAGAGGGAATGTAGCAACAGAGGCTCCTAGGC 1813
Db      1849 AGC-----TTGCGAAGACAGCAGTGTCTGTGGCAGCTACAGGCGCATCAGAGT 1896
Qy      1814 CCCATCTCCATTCTTAGGTTAAGAGAGCATTTCTCAGACTCCAGGCGGAGACTGA 1873
Db      1897 GACAAGTAATC-----AACCACCTCCTCTTTCACTGA 1927
Qy      1874 GCCTAGCTTCAGCAACCAAGGTTCTCTGGGACCCAAAGTTTATGGGAGAGGCAAG 1933
Db      1928 ACCAGCCTTAGGAGTTAGGTTTCCCGAGCTCTGGACAAAGGCTTAAGACAAGG 1987
Qy      1934 ACTTCATGGGAAGAGAGGAGGAGGCGCTGGGTAGAAACGCTGTGGTGTCTCTTTGG 1993
Db      1988 AAGAGAAGGTGGGTA-----TGATGGAATACTTGTGTGCTGTCTTGCCTTC 2034
Qy      1994 CTTTAAAGCAAAAGCGTCATCTTGGCCTCTACCTCTGATAGCTTGAAGGTTGGCAA 2053
Db      2035 CAG-----GACTCCCTCCCTGCTTCCCTTATTTCTTAAGTCGCGCTTGGCAA 2082
Qy      2054 CCACACTGTGCTACAGGTGAGGAGGAGGAGCTCTTCCCTCAGAGTGTCTGTACAG 2113
Db      2083 CTGCTGTGTGCTACCAAGGAGGAGGAGAGACACCTCTCCAGAGTCTATGCTCAGG 2142
Qy      2114 AAGTTCTTTAAACCCATATGCCCCAAGATAGCTGTGAGGAGGCGCTTTAAAGACGAA 2173
Db      2143 AAGTTCTTTAAACC--TGTGTTCAAGAGTAGCTTGAAGGAGGCGCTGAAAAAACA 2200
Qy      2174 CAAGTAATTTA----COAGTTCTACTGGGTTCTTCCCAACCTCCCAAGTGGCGAG 2228
Db      2201 AAACAACAAACAAAGCCACCTGTGAGGTTTATCTTACATTCCTCAAGTGGGAGGA 2260
Qy      2229 GCCTAAGAGAGGCTATTTCTTAAGCCACACATAGCTGCACCTCGTGGCTGCAGCCAAA 2288
Db      2261 GCTCAGGTAAGAGAAGAGTCCCAAGATGCATTTCAAGCCTGGGCTGGAGCCAAAGAGA 2320
Qy      2289 ACAAGAACTGGGTGT---TGAGTATTCATCACTAAGAACCC---AAATCCAGGCACT 2342
Db      2321 GGAAGAACTGACACTGACCGAGCATTCAGAGCAAGGTCCTTTCAATCCAGGGCACA 2380
Qy      2343 CATATGTGAAGTAAGAAGCTCACTTCTTACTCTCTCCAAA--GAAGTGGGAAGA 2400
Db      2381 GCTGTGATGACAGAGAGCCCAACCTCTTTAGTCTCTTAAAGAGGAGATGGGGTGAA 2440
Qy      2401 ACCATCAAACTTCTCCTGCTGACTTACCAACACAGGAAACAGCAGGAGGAGGCTCA 2460
Db      2441 ATCAATGGGATTTCTCCTGCTGACTAGCAACAGGAGGAGGAGGAGCAAGAGCTCA 2500
Qy      2461 GGACTTA 2467
Db      2501 GGATTTA 2507

```

RESULT 9

RATRAB15X

LOCUS

DEFINITION

ACCESSION

VERSION

945 bp mRNA linear ROD 27-APR-1993

RATRAB15X

Sprague-Dawley (clone LRB9)

RAB15 mRNA, complete cds.

M83679.1

GI:206536

```

KEYWORDS  LMW GTP-binding protein.
SOURCE     Rattus norvegicus (strain Sprague-Dawley) (library: LAMBDA ZAPII)
           adult brain cDNA to mRNA.
ORGANISM  Rattus norvegicus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
           Rattus.
REFERENCE 1 (bases 1 to 945)
AUTHORS   Elferink,L.A., Anzai,K. and Scheller,R.H.
TITLE      rab15, a novel low molecular weight GTP-binding protein
           specifically expressed in rat brain
JOURNAL   J. Biol. Chem. 267 (9), 5768-5775 (1992)
MEDLINE   92210533
PUBMED    1313420
FEATURES   Location/Qualifiers
            1..945
            /organism="Rattus norvegicus"
            /strain="Sprague-Dawley"
            /db_xref="taxon:10116"
            /tissue_type="brain"
            /dev_stage="adult"
            /tissue_lib="LAMBDA ZAPII"
            1..945
            /gene="RAB15"
            220..858
            /gene="RAB15"
            /codon_start=1
            /product="RAB15"
            /protein_id="AAA41995.1"
            /db_xref="GI:206537"
            /translation="MAKQYDVLFRLLIGDSVGKTCLLCRFTNEFHSHSTIGVD
            FMKTIIVDGIKVIQIWDTAGRYQTITKQYRAAGIQLVDYISSEYQHIMKW
            VSDVEYAPGVQKILIGNKADEQKQVREGQQLAKYGMDFYTSACTNLNIKE
            SFTRLTELVLQAHKRLDGLRTCASNELALAELEDEKTEGPANSSKTCWC"
BASE COUNT 222 a 259 c 292 g 172 t
ORIGIN
Query Match 17.7%; Score 577.6; DB 10; Length 945;
Best Local Similarity 88.7%; Pred. NO. 1e-137;
Matches 637; Conservative 0; Mismatches 79; Indels 2; Gaps 1;
Qy      21 TTCCGCGCCCGCTGGCCCGAGTATGCGGAGCAGTACGATGCTCTTCGGGTGCTG 80
Db      196 TTCGGGCCCCCTTGAGTCCCTGCCATGGCGAAACAGTACGATGTCTATTCGGGTACTG 255
Qy      81 CTGATCGGGGACTCGGGGTGGCAAGACCTGCTGCTGTCGCGCTTCACGACACAGAG 140
Db      256 CTGATCGGGGACTCGGGGTGGCAAGACCTGCTGCTGTCGCGCTTCACGACACAGAG 315
Qy      141 TTCACCTCTCGCACATCTCCACCATCGGTGTTGACTTTAAGATGAAGACCATAGAGTA 200
Db      316 TTCACCTCTCGCATATCTCCACCATCGGTGTTGACTTTAAGATGAAGACCATCGAAGTA 375
Qy      201 GACGSCATCAAGTCCGGATACAGATCGGACACATCGCAGGCGAGGAGATACAGACC 260
Db      376 GACGSCATCAAGTGGAGTACAGATTTGGGACACAGCAGGCGAGGAGGTACCACT 435
Qy      261 ATCAAAAGCAGTACTATCGGCGGCCCGAGGGGATATTTTGGTCTATGACATTAAGCAGC 320
Db      436 ATCAAAAGCAGTACTATCGGCGGCCCGAGGGGATATTTTGGTCTATGACATTAAGCAGC 495
Qy      321 GAGGCTCTTACCAGCACATCATGAAGTGGGTGAGTGACGTGGATGATGATGACGACAGAA 380
Db      496 GAGGCTCTTATCAGCACATCATGAAGTGGGTGAGTGACGTGGATGATGATGACGACAGAA 555
Qy      381 GCGTCCAGAGATCCTTATTGGGAATAAGCTGTATGAGGACCAACACGCGAGTGGGA 440
Db      556 GCGTCCAGAGATCCTTATCGGGAATAAGCTGTATGAGGACCAACACGCGAGTGGGG 615
Qy      441 AGAGAGCAGGCGAGCAGCTGGCGAAGGAGTATGGCATGGACTTCTATGAACAAGTGC 500
Db      616 AGAGAGCAGGCGAGCAGCTGGCTAAGAGTATGGCATGGACTTCTATGAACAAGTGC 675

```


AC096084 162495 bp DNA linear HTG 24-AUG-2002
 Rattus norvegicus clone CH230-16122, *** SEQUENCING IN PROGRESS
 ***, 70 unordered pieces.

AC096084
 AC096084.6 GI:22164942
 HTG: HTGS-PHASE1.
 Norway rat.
 Rattus norvegicus

ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE
 1 (bases 1 to 162495)
 Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Blewett, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buha, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, J., Ceasar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escoto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C. M., Gabist, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gregeorge, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hayes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorensu, L., Loulseg, H., Lozado, R. J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mawhinney, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A.,
 Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K.,
 Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D.,
 Newton, N., Nguyen, N., Norris, S., Nwaokeme, O., Okwuonu, G.,
 Olarpunagoo, A., Pal, S., Parks, K., Pasternak, S., Paul, H.,
 Perez, A., Perez, L., Plankoch, C., Plopper, F., Polindexter, A.,
 Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E.,
 Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y.,
 Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A.,
 Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S.,
 Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A.,
 Sisson, I., Sitter, C. D., Smales, D., Sneed, A., Sodergren, E.,
 Song, X.-Z., Sorrell, R., Sosa, J., Steimle, M., Strong, R., Sutton, A.,
 Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S.,
 Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D.,
 Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J.,
 Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R.,
 Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S.,
 Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X.,
 Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R.,
 Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Direct Submission
 Unpublished

REFERENCE
 2 (bases 1 to 162495)
 Worley, K. C.

AUTHORS
 Direct Submission

JOURNAL
 Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

REFERENCE
 3 (bases 1 to 162495)
 Rat Genome Sequencing Consortium.

AUTHORS
 Direct Submission

JOURNAL
 Submitted (24-AUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One

1247: contig of 1247 bp in length
 1347: gap of unknown length
 2634: contig of 1287 bp in length
 2734: gap of unknown length
 3987: contig of 1253 bp in length
 4087: gap of unknown length
 5288: contig of 1201 bp in length
 5388: gap of unknown length
 5588: contig of 1198 bp in length
 6686: gap of unknown length
 8143: contig of 1457 bp in length
 8243: gap of unknown length
 9249: contig of 1006 bp in length
 9349: gap of unknown length
 10754: contig of 1405 bp in length
 10854: gap of unknown length
 12161: contig of 1307 bp in length
 12261: gap of unknown length
 13639: contig of 1378 bp in length
 13739: gap of unknown length
 14946: contig of 1207 bp in length
 15046: gap of unknown length
 16342: contig of 1296 bp in length
 16442: gap of unknown length
 17772: contig of 1330 bp in length
 17872: gap of unknown length
 19007: contig of 1135 bp in length
 19107: gap of unknown length
 20796: contig of 1689 bp in length
 20896: gap of unknown length
 22145: contig of 1249 bp in length
 22245: gap of unknown length
 24330: contig of 2085 bp in length
 24430: gap of unknown length
 25589: contig of 1159 bp in length
 25689: gap of unknown length
 27349: contig of 1660 bp in length
 27449: gap of unknown length
 28577: contig of 1128 bp in length
 28677: gap of unknown length
 30291: contig of 1614 bp in length
 30391: gap of unknown length
 32126: contig of 1735 bp in length
 32226: gap of unknown length
 33343: contig of 1117 bp in length

1247: contig of 1247 bp in length
 1347: gap of unknown length
 2634: contig of 1287 bp in length
 2734: gap of unknown length
 3987: contig of 1253 bp in length
 4087: gap of unknown length
 5288: contig of 1201 bp in length
 5388: gap of unknown length
 5588: contig of 1198 bp in length
 6686: gap of unknown length
 8143: contig of 1457 bp in length
 8243: gap of unknown length
 9249: contig of 1006 bp in length
 9349: gap of unknown length
 10754: contig of 1405 bp in length
 10854: gap of unknown length
 12161: contig of 1307 bp in length
 12261: gap of unknown length
 13639: contig of 1378 bp in length
 13739: gap of unknown length
 14946: contig of 1207 bp in length
 15046: gap of unknown length
 16342: contig of 1296 bp in length
 16442: gap of unknown length
 17772: contig of 1330 bp in length
 17872: gap of unknown length
 19007: contig of 1135 bp in length
 19107: gap of unknown length
 20796: contig of 1689 bp in length
 20896: gap of unknown length
 22145: contig of 1249 bp in length
 22245: gap of unknown length
 24330: contig of 2085 bp in length
 24430: gap of unknown length
 25589: contig of 1159 bp in length
 25689: gap of unknown length
 27349: contig of 1660 bp in length
 27449: gap of unknown length
 28577: contig of 1128 bp in length
 28677: gap of unknown length
 30291: contig of 1614 bp in length
 30391: gap of unknown length
 32126: contig of 1735 bp in length
 32226: gap of unknown length
 33343: contig of 1117 bp in length

NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 NOTE: This is a 'working draft' sequence. It currently
 consists of 70 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 Project Information
 Center project name: GEJZ
 Center clone name: CH230-16122
 Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 116551 bases at least Q40
 Consensus quality: 123470 bases at least Q30
 Consensus quality: 127758 bases at least Q20

On Aug 9, 2002 this sequence version replaced gi:21723229.
 Baylor Plaza, Houston, TX 77030, USA

* 33444	33443: gap of unknown length	125410	CCCTCCCCAGTCCCTTCCCTCAGGAGGCCGCTGGAGCA---GGAGTCTGGGCTTG	125354
* 34474	34975: contig of 1532 bp in length	Qy	CCCTGCTGCTGCTCTCTCGTGTGATGACCTATTAGTATCAGTAGCCTACTCTCCCT	820
* 34976	35075: gap of unknown length	Db		
* 35076	37579: contig of 2504 bp in length	Db	TCCATC-ACGTGCTCTTTCATTTGAGATCTGTTGAAACACAGTAGTCTACT-CCCT	125296
* 37580	37679: gap of unknown length	Qy	821 GCCTGGCCCTGAGAGCGGCTGCTGCTATCT-CAAGCAGCCCTGTCGCCAGCCGCTCC	879
* 37680	39003: contig of 1324 bp in length	Db		
* 39004	39103: gap of unknown length	Db	125295 GCCTGGCCCTGAGCAGCTTGTCTGGTCTCGAGCAACCTCTGTCGCCAGCCCTGCT	125236
* 39104	40564: contig of 1461 bp in length	Qy	880 ACCCTGAGTGGTCTTCTCAGCCTGTTTCCAGCCACAGCCGCTGCTACGACCCCAAG	939
* 40665	40664: gap of unknown length	Db		
* 40665	41979: contig of 1315 bp in length	Db	125235 ATCCTGGAGTGGTAACCTTTACCCCTGTTTCCAGCCGAGGCTGCTGTGACCCC	125180
* 41980	42079: gap of unknown length	Qy	940 ATGTGCGCGAGCAGCTGTCTCACCATCCCGCACCCACAGACACAGCCAGGCTGGAGT	999
* 42080	43598: contig of 1519 bp in length	Db	125179 AGAATGTGCTTCGACCATCTCACCAACCCAGGCTCTGGGACACGCTCAAGGCTGGAGT	125120
* 43599	43698: gap of unknown length	Qy	1000 CCAGGCACCTTTCAGCTGCTCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1059
* 43699	45418: contig of 1720 bp in length	Db	125119 CAAAGCCACTTAAAGACTCCCTTTAGTGCATCTCGTGGGCTCTCTGCTTTCTCTCT	125060
* 45419	45518: gap of unknown length	Qy	1060 CTTCCTTCTTCTTCTTCTGACCTTCCCTCCGCTGGTTCGTATCAAAAGCTCT	1119
* 45519	47359: contig of 1841 bp in length	Db	125059 CTTCCTTCTTCCAGTCCCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	125000
* 47360	47459: gap of unknown length	Qy	1120 CAAACCCGCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1179
* 47460	48918: contig of 1459 bp in length	Db	124999 C---CTCCCTAGCCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	124943
* 48919	50383: contig of 1365 bp in length	Qy	1180 TATCAAGGGGATGAGCCAGGCTGCTGGGGAGGTTCCACCTTGGATCCAGGAAGAAC	1239
* 49019	50384: gap of unknown length	Db	124942 TGCCTAAGGAATGACCCAGGCTC-----TGCCCCAGGGAGGATG	124902
* 50384	50483: contig of 1863 bp in length	Qy	1240 CTCACCTGCTGCTGCTGGTGGGCGCAAGGCTACAGGCTGCTTCTCTTCCCTCCAC	1299
* 52347	52446: gap of unknown length	Db	124901 GCTCACCTG-CTCTGGGCTGGGCTGAAGGCTA-AGGCTGCTTCTTCTCTCTCC	124846
* 52447	52557: contig of 2811 bp in length	Qy	1300 CCCACTGCTCCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1359
* 52558	53537: gap of unknown length	Db	124845 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	124787
* 53538	57005: contig of 1648 bp in length	Qy	1360 GAGGTAGGAGGAAACAAACCGGGAGTCTCTGAGCTGCTGGGCTGCTGCTGCTGCTGCT	1419
* 57006	57105: gap of unknown length	Db	124786 GAGGTAGGAGGAAACAACTATGAGTG-GTCTCAAGCT-GGACTGCTGCTTCTGCT	124729
* 57106	59670: gap of unknown length	Qy	1420 CATTCGCCAGCAGAGCTTTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1479
* 59671	61952: contig of 2182 bp in length	Db	124728 TTGCTTCTGCTGGAGTTTGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	124669
* 61953	62052: gap of unknown length	Qy	1480 GCTCAGAGGAGGCTGCTTCAAGAGGAAACAAATGAGGGTGGCAGGATAAAAGTC	1539
* 62053	64503: contig of 2451 bp in length	Db	124668 GTCCAGAGGAGGACACAGCAGAGAAGAGTACAAAGATGGGTTGGCGGGATTAAGTCA	124609
* 64504	65752: contig of 1149 bp in length	Qy	1540 ACCTCCATCTCTACTCCCATGACATGAACAATTTCTCTCACCTGGCTCCCAAA	1599
* 65753	65852: gap of unknown length	Db	124608 --CTCTGTTCTCTACTCTCCCATGACGAGTGACACA--GCCCTTTCACCGGGTTCG	124553
* 65853	68166: contig of 2314 bp in length	Qy	1600 TTAAAGATGAGGACCAAGGCTGTTGGTACTCCAGGGCAGAGGAGCCCTGGGGTTCAG	1659
* 68167	68466: gap of unknown length	Db	124552 ATTTAAAGATGAGGATCAAGGCTGAGGACGC-CAGAGGGCAAGGGGATCCCGAGGTC	124495
* 68267	70091: contig of 1825 bp in length	Qy	1660 TGACACTGTCCAGGCCAACCTCACTCCCAAGGGGAGCATTTTGAATGAAGACTAG	1719
* 70092	72887: gap of unknown length	Db	124494 -GACACTGTGTAGTCACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG	124436
* 72888	72987: gap of unknown length	Qy	1720 CTCCTATGTATCAGGTTAAGACCAAGG-AGAGCTGGCCAGGACAGCAGTTTGCACAGC	1778
* 72988	74782: contig of 1795 bp in length	Db	124435 CTCCTAGATGCGAGGTTAGGACAAAGATAGAGGGGCTTACAGAACAGTGTGCTGT	124376
* 74783	77248: contig of 2366 bp in length	Qy	1779 AGAGGGATGTAGCAACAGCAGGCGCTCTAGGCCCCCATCTTCCATTCTTAGTGAAGA	1838
* 77249	79379: gap of unknown length	Db	124375 GGTGACAGGAGGCGGTCCTCCCTGTGAGGACATTTACCAGAGACTCCTCGCAGAGGT	124316
* 79380	83158: contig of 3779 bp in length			
* 83159	83259: gap of unknown length			
* 83259	85363: contig of 2105 bp in length			
* 85364	85463: gap of unknown length			
* 85464	88527: contig of 3084 bp in length			
* 88528	88627: gap of unknown length			
* 88628	90451: contig of 1824 bp in length			
Query Match 13.0%; Score 423.6; DB 2; Length 162495;				
Best Local Similarity 61.8%; Pred. No. 1.5e-97; Mismatches 664; Indels 95; Gaps 31;				
Matches 1227; Conservative 0;				
Qy	523 AGTCATTCAGCGCTGACAGAGCTGGTCTGCTGAGGCGCCATAGGAGGAGCTGGAGGCC	582		
Db	125590 AGTCCTTCACTCGTCTGACCGAGCTGCTGCTGAGGCTCACAGAAAGAGCTGGATGGTC	125531		
Qy	583 TCCGGATGCTGCCAGCAATGAGTTGGCACTGGCAGAGCTGAGGAGGAGGAGGCAAC	642		
Db	125530 TCCGAACATGTGCCAGCAATGAGTCCGACTGGCCGAGCTGGAGGAGGAGGCAACAA	125471		
Qy	643 CCAGAGGCCAGGAACTCTTCCGAAACCTGCTGGTGTGA--GTCTGTGTGGGCAAC	700		
Db	125470 CTGAAGGCCAGCAACTCTTCAAGACCTGCTGGTGTGAGGCTCTCTGTGTGAATCT	125411		
Qy	701 CCACAGCACACCTCTTCCCTCAGGAGGCCGCTGGGACACAGGAGGAGCCGGGCTTG	760		

```
QY 1839 AGAGCATTTCTCAGACTCCAGCGGAGGACTGAGCCTTAGCCTTCAGCAACCAAGGTTT 1898
D 124315 AGTGATATAGTCAA---TCACTCCTCGTTCACTGAACCCAGCCTTCGGAAGTTAAGGTTT 124259
QY 1899 TCCTGGGACCAAGTTTATGGGAGAGGCAAGACTTCATGGGAGAGAGAGGAGG 1958
D 124258 TCC-----CCAGAGTCCTGGACACAGCCCT---TAAACATGCCAGAGAGGCGG 124208
QY 1959 CCTGGGTAGAAACGCTGGTGTCTCTTTGGCCTTTAAGACAAAGCGCTCATCTTG 2018
D 124207 GTATGGGTAGAATA-CTTGGTGTCTTTCGCTTCAGGACTCC---TCCCTTGTTC 124152
QY 2019 CCTCTACCTCCGTATAGGCTTCAGGTTTGCCAAACCACTGTGCTACAGGTGAGGG 2078
D 124151 CTTATTCTTAAGTCTAAGCTGAAGCTTTGCCAACTGCTGAGTGTCTACTCAGGAGGG 124092
QY 2079 AAGAGGACTCCTTCCCTCCAGAGTGTATGTTTCAGGAAGTTTCTTAACCCCATATGCCCC 2138
D 124091 AAGGACA---CCTCTCCAGAGTCTATGCTCAGGAAGTTTCTTAACCC--TGTGCCCC 124037
QY 2139 AAGAGTAGCTCGTAGGAGGCCCTT-----TAAAGACGGGAACAAAGTAATTTACAGT 2189
D 124036 AAGAGTAGCTTGAAGGAGGCCCTCAGAAAAAAGAAAGAAAGCAAAACCAAC 123977
QY 2190 TCTACTGGGTTCTCTGCCACCGTCCAGAGTGGGAGGCCCTAGGAAGAGGTCATCT 2249
D 123976 CTTGGTAGGTTTGTCTTACATTCCTCAAGTGGGAGGCTCAGGTAAGAGAATAGTCC 123917
QY 2250 TAAGCCACACATTAGCTGCTCGCTGGCTGCGAGCCAAACAAAGAACTGGGTGTTGAGT 2309
D 123916 CAAGATGCAC-TTCAAGCTGGGCTGGAGCCAAACAAAGGAGGAGCTGGGCTCTGATG 123858
QY 2310 ATTCACTCACTAAGAACCA-----AAATCCAGGCGCACTATATGTAAGGATAA 2358
D 123857 AGCATTCACTGACTGACCGAGGATTTTCAATCCAGGCGCACGCGTGATCAGAGA 123798
QY 2359 GAACCTCACTTCTTACTCTCAAAAGAG--TGGGGAAGAACCACTCAACCTTCTCT 2417
D 123797 GCCTCACACCTTTTATGCTCTTAAAGAGGAGATGGGATGAAGCAATGGGAATCTCT 123738
QY 2418 CTTGACTTACCAACACAGGA---AAACAGCAGGAGGAGGCTCAGGACTTAGGACAG 2474
D 123737 CTTGACTTACCAACAGGAGAGGATTCAGCGGGAGAGACAGCTCTGGATCTGTTGGA 123678
QY 2475 GGTATA 2480
D 123677 GGGAAA 123672
```

RESULT 13

```
BC013790
LOCUS BC013790
DEFINITION Mus musculus, clone MGC:6897 IMAGE:2655151, mRNA, complete cds.
ACCESSION BC013790
VERSION BC013790.1 GI:15489393
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1. (bases 1 to 2560)
Strausberg,R.
Direct Submission
Submitted (04-SEP-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
```

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: ang@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 5 Row: 1 Column: 20
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, GenesScan gene prediction.

FEATURES

source	Location/Qualifiers
1..2560	/organism="Mus musculus"
	/db_xref="taxon:10090"
	/map="FVB/N-3"
	/clone="MGC:6897 IMAGE:2655151"
	/tissue_type="Mammary tumor. MMTV-LTR/INT3 model. 5 month
	old mouse. Taken by biopsy."
	/clone_lib="NCI-CGAP_Mam2"
	/lab_host="DH10B"
	/note="Vector: pCMV-SPORT6"
154..660	/codon_start=1
	/product="Unknown (protein for MGC:6897)"
	/protein_id="AAH13790.1"
	/db_xref="GI:15489394"

CDS

BASE COUNT	614 a	648 c	719 g	579 t
ORIGIN				
Query Match	12.8%	Score 416.4	DB 10	Length 2560
Best Local Similarity	91.5%	Pred. No. 5	1e-96	
Matches	441	Conservative	0	Mismatches 41; Indels 0; Gaps 0;

```
QY 22 TCCCGGCCCGCTGGCCCGCAGTCAATGGCGAAGCAGTACGATGCTCTTCGGGCTGCTGC 81
D 131 TCGGGCCCGCTTGAAGTCCCTGCCATGGCGAAGCAGTACGATGCTCTTCGGGCTGCTGC 190
QY 82 TGATCGGGGACTCCGGGTTGGCAAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 141
D 191 TGATCGGGGACTCCGGGTTGGCAAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 250
QY 142 TCCACTCTCTCGACATCTCCACCATCGTGTGTTGACTTTAAGATGAAGACCATAGAGGTAG 201
D 251 TCCACTCTCTCGACATCTCCACCATCGTGTGTTGACTTTAAGATGAAGACCATAGAGGTAG 310
QY 202 ACCGCATCAAGTGGGATACATCTGGGACACTCGGAGGAGGAGAGATACAGACCA 261
D 311 ACCGCATCAAGTGGGATACATCTGGGACACTCGGAGGAGGAGAGATACAGACCA 370
QY 262 TCACAAGCAGTACTATCGCGGGCCCGAGGGATATTTTGGTCTATGACATTAGCAGCG 321
D 371 TCACAAGCAGTACTATCGCGGGCCCGAGGGATATTTTGGTCTATGACATTAGCAGCG 430
QY 322 AGCGCTCTTACCAGCATCATGAAGTGGGTGAGTGCAGTGGGATGAGTACGACCAAG 381
D 431 AGCGCTCTTACCAGCATCATGAAGTGGGTGAGTGCAGTGGGATGAGTACGACCAAG 490
QY 382 GCGTCAGAGAGATCTCTATTGGGAATAAGGCTGATGAGGAGCAGAAACCGCAGGTGGAA 441
D 491 GAGTCCAGAGAGATCTCTATTGGGAATAAGGCTGATGAGGAGCAGAAACCGCAGGTGGGA 550
QY 442 GAGAGCAAGGCGCAGCAGCTGGCGAGGAGTATGGCAGTGGACTTCTATGAACAAGTGCCT 501
```


Db	135269	CTGAGGCCAGCAAACTCTTCAAGACCTGCTGGTGTGAGGTTCTGAGTGGAGTCCC	135211
Qy	703	ACAGCACACCCCTCTTCCCTCAGAGAGCCCGTGGGCAGACAGGGAGCGGGGCTTTGCC	762
Db	135209	CC---CAGCCCTTCCCTCAGAGAGCCCGTGGAGCA---GGAGTCTGGGCTTTGCC	135157
Qy	763	CTGCTGCTGCTCTCTGCTGATGACCCCTATTGAGTATCAGTAGCCACTACTCCCCCTGC	822
Db	135156	C-ACGACCGCTCTTTCATTTGACAACTCTGTTGAAAACCACTAGCTGCTACT-CCCTGC	135099
Qy	823	CTGGCCCTGAGAGCGGCTGCTGCTCATCTC-AGCAGCCCTCTGCCAGCCGTC-CA	880
Db	135098	CTGGCCCTGGAGCAGCTCTGCTGGTCTCTGACAGCCCTCTGCCCCAGCCCTCTAA	135039
Qy	881	CCCTGGAGTGTCTTCTTCAAGCCTGTTTCCCAGCCACAGCCCTGCTACGACCCCCAGCA	940
Db	135038	TTCTGGAGTGTGACCTTTTCAAGCCTGTTTCCACGCGCAGGCCCTCTCTGA-CCCGCAA	134980
Qy	941	TGTGCCGGAAGCACTGTCTACCATCCGGACCCACACAGCAACAGCAGGCGTGGAGTC	1000
Db	134979	TGTGCTGC-----GACCACCCAGGCTCTCTGGGAGCAGGCTCAAGCTGGAGTC	134931
Qy	1001	CAGGCCACTTTACGCTGCTCTTCTTCGCTGCATCTGTCATCTCTGCTGCTTTTCTCTC	1060
Db	134930	AAGCCCACTTCAAG-----ATTCTTAGTGCACTCGTGGGCTCTCTGCTTTTCTCT-	134878
Qy	1061	TTCCCCCACTTCTTCTCTGACCCCTCCCTCCGGTGGCTTTCTGATCAAGACTCCTC	1120
Db	134877	--TCTTCTCTCTTCTTCTCACCTCCCTCCCTTGTAAATTCGCATCCCTCATCATC	134821
Qy	1121	AAACCCCGTCCCGGTGTCTGCTGTGTGCAGCTGCTCTTCTTCTTCTTCTTAAGCT	1180
Db	134820	CCAGTCTCTCTCTCTGCTGTGTCTGCTGCCAAGGCTGCTACTGCACTTCTCTCTT	134761
Qy	1181	ATCCAAGGGATGACACCAAGCTCGTGGGAGGTTCCACCTTGGATCCAGAGAAACCC	1240
Db	134760	TTTT--CCTTCTGATTTGCCTAAGGAATGGACCCAGGCTCTGCCCCAGGAGGACTG	134703
Qy	1241	TCCACCTGCTCTGGTGGGCGCAAGGCTACAGGCTGCTTCTTCTTCTTCTTCCCCACC	1300
Db	134702	TTACACCTGTGTCG-GGGTGGCTGAAGGCTA-AGGTCCTTCTTCTTCTTCTTCTTCT	134654
Qy	1301	CCACTGTCCTCATGTGCGATGGCCCTGCTTCCCGAGTACCTCGCAAGTGGAGCATCG	1360
Db	134653	-----TGTGCGCTGGG-CTGCTTCCCGAGTGGCTGGGAGGTAGAGCACTG	134608
Qy	1361	AGGTAGAGGGAACACAAACCGGGAGTCTCGAGCTGGGGCTGGCCTACCTTACCTACC	1420
Db	134607	AGGTAGAGGAACACTA--CTAGTGGGCCCTCAGCCTGGACTGGGCTCTTCTGCGCTT	134550
Qy	1421	ATTTCCCGACACAGAGCTTTGCGCTTGTGTGGCTGCCCGCCCTGCTTCTTGGGGAACTGAG	1480
Db	134549	TGCCTTCTGCTGACGTTCTGGCCCTTGGCTGGCTGCCCTGCTGTGCGCTCTGTAGGAAC	134490
Qy	1481	CTCAGAGCAGGTGCTCTCAGAGAAGGAACAAATGA-----GGGTGGCAGGGATAAA	1534
Db	134489	TCCAGAGGCAGGCACAGCAGAGAAGGAAAACGAAGAACAACGATGGGTGGCAGGAT--T	134432
Qy	1535	AAGTCACCTCCATTTCTACTCTCCATCAGCAGCATGAACAAATTTCTCTCCACCTGGCTC	1594
Db	134431	AAGTCACCTCTGTTCTTACCTCCATGCCCTC-----TCTCTTACCTTGGATTC	134383
Qy	1595	CCAAATTTAAGATGTGGCAAGGCCCTGTGGGTACTCCAGGGGCAAGGAGAGCCCTGGG	1654
Db	134382	ACAAATTTAAGATGGGATCGAGGCTGTAGGACAC--CAAGGCAAGGGAGCCCTGGG	134324
Qy	1655	GTCAGTGACACTGTGAGGCCCAACCATGCATCCCAAAAGGGAGGATTTTGGAAATGAAG	1714
Db	134323	GTCA--GATACTGTATAGTCAGCTGTGAATCCAGATAGAGCAGAGCGGATGCGAAGG	134266
Qy	1715	ACTAGTCTCTATGTATCAGGTTTAAGACGAAGGGAGAGCTGCCAGGACAGCATTTGCA	1774

Db	134265	CGGGCTCCTAGAACACCGAGGGTTAGAGCAAGACGAGAGGAGC-----TTCGCA	1342121
QY	1775	CAGCAGAGGGGAATGTAGCAACACGACGAGGCGCTCCTAGGCCCCATCTTCCATTTCTTAGST	1834
Db	134217	GAAACGNGTGTGTGCGGACTACAGGGCATCAGAGTGCACAGTAATC-----	134168
QY	1835	AAGAAGAGCATTTCTCAGACTCCAGCGGAGGACTGAGCCTTACAGCAACCAAG	1894
Db	134167	-----AACCACCTCCTTTCACTGAACCCAGCCCTAGGAGGTAGG	134127
QY	1895	GTTCTCTGGGACCCAAAGTTTATGGGAGAGGGCAAGACTTCATGGGAGAGAGAGG	1954
Db	134126	GTTTCCCAGAGTCTGGACAACAGGCCCTTAAAGCAAGGGAAGAGAGGTTGGTA-----	134072
QY	1955	AAGGCCCTGGGTAGAAACGCTTTGGTGTCTCTCTTTGGCCTTTAAGACAAAGCCTCAT	2014
Db	134071	-----TGATGGAATACTTGGTGTCTTGCCTTCC-----AGGACTCC	134032
QY	2015	CTTGCCCTTACCTTCTCATAGGCTTTAGGGTTTGGCCAAACACACTGTGGCTACAGGTGG	2074
Db	134031	CTCCCCCTGGTTCCTTATTTCTAAGTGGGCTTTGGCAACTGCTGTGTGGCTACCAAGG	133972
QY	2075	AGGGAAGAGGACTCCTTCTCCAGAGTCTATGTTTCAGGAGTTTCTTTAAACCCATATG	2134
Db	133971	AGGGAAGAGACACCTCCTCCAGATTTCTATGCTCAGGAAGTTTCTTTAAACCC--TG	133914
QY	2135	GCCCAAGAGTAGCTCGTAGGAGGCCCTTTAAAGACGGGAACAGTAATTTA-----C	2185
Db	133913	GTTCAAGAGTAGCTTGAAGGAGGCCCTGAAACAACAACAACAACAACAACAAGC	133854
QY	2186	CAGTTCTACTGGGTTCTGCGCCACCGTCCCAAGGTGGCGAGGCCTTAGGAGAGGGTCA	2245
Db	133853	CAACCTGGTAGGGTTTATCTTACATTTCCCAAGGTGGGAGGAGCTCAGGTAAGAAGA	133794
QY	2246	TTCTTAAGCCACACATTAGTGTGCTGCTGCATGCCCAACCAACAAGAACTGGGTGT-	2304
Db	133793	GTCCCAAGATCGACTTCAAGCCTGGGGCTGGAGCCAAAGCAGAGAAGAACTAGGCATG	133734
QY	2305	--TGAGTATTTCATCAACTAAGAAC--AAATCCAGGGCACTCATATGTGAAGGATAG	2359
Db	133733	ACCGAGCATTCAGGACCAAGTCTCTTCAATCCAGGCACAGCTGTGATCACAGAGAG	133674
QY	2360	AACCTCACTTCTTACTCTCTCCAAAA--GAAGTGGGGAAGAACCATCAACCTTTCT	2417
Db	133673	CCCCACACCTTTAGTCTCTTAAAAAGGAGATGGGGGTGAAATCAATGGGGAATTC	133614
QY	2418	CCTGACTTACCAAAACAGGAAACAGCAGGAGGGTGGCTCAGGACTTA	2467
Db	133613	CCTGACTTACCAAAACAGGAAACAGCAGGAGGAGAACAGCTCAGGATTTA	133564
RESULT 15			
AX341683		313 bp	DNA
LOCUS	AX341683		linear
DEFINITION	Sequence 1930 from Patent WO0196388.		PAT 10-JAN-2002
ACCESSION	AX341683		
VERSION	AX341683.1	GI:18137665	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1. Jiang Y., Harlocker S.L. and Secrist H.		
JOURNAL	Compositions and methods for the therapy and diagnosis of colon		
FEATURES	cancer		
	Patient: WO 0196388-A 1930 20-DEC-2001;		
	CORIXA CORPORATION (US)		
	Location/Qualifiers		
	1...313		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	68 a 85 c 85 g 73 t		2 others

ORIGIN

Query Match 9.4%; Score 307; DB 6; Length 313;
Best Local Similarity 99.0%; Pred. No. 6.1e-68;
Matches 308; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY	2824	ACCTTAGGGTTCGGAGGGAACATGGAAATCAGGATTCATAGACTGATAGCCCTA	2883
DB	1	ACCTTAGGGTTCGGAGGGAACATGGAAATCAGGATTCATAGACTGATAGCCCTA	60
QY	2884	TCCACAAGGCCATCAGTGGGAAGGTATCGGAGCAGAGAGAAATGGGATTTAGGG	2943
DB	61	TCCACAAGGCCATCAGTGGGAAGGTATCGGAGCAGAGAGAAATGGGATTTAGGG	120
QY	2944	TGCAGCTACGCTCACCCCTAAACTTTTGGTGGCCCTGGGCGATGCTTTGAGGCCCAGACTGT	3003
DB	121	TGCAGCTACGCTCACCCCTAAACTTTTGGTGGCCCTGGGCGATGCTTTGAGGCCCAGACTGT	180
QY	3004	TAAGCAGGCTCTGCTGGCCCTGTTTACTCGTCACCACTCTGCACCTGCTGCTTGAGACT	3063
DB	181	TAACCAAGGCTCTGCTGGCCCTGTTTACTCGTCACCACTCTGCACCTGCTGCTTGAGACT	240
QY	3064	CCATCCAGCCGAGCAGCCAGCTGCTCCAGCCTCCACTATCTCCCTGTCAGCGGTG	3123
DB	241	CCATCCAGCCGAGCAGCCAGCTGCTCCAGCCTCCACTATCTCCCTGTCAGCGGTG	300
QY	3124	AACTTCGTGTA	3134
DB	301	AACTTCGTGTD	311

Search completed: June 23, 2003, 15:56:06
Job time : 8188 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: June 23, 2003, 11:52:42 ; Search time 668 Seconds
(without alignments)
10980.177 Million cells

Title: US-09-817-198B-1

Perfect score: 3257
Sequence: 1 tgcctgtgtccgcgcgcag.....aaaaaaaaaaaaaaaaaaaaa 3257

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*

- 1: /SID\$2/cgcdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SID\$2/cgcdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SID\$2/cgcdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SID\$2/cgcdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SID\$2/cgcdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SID\$2/cgcdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SID\$2/cgcdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SID\$2/cgcdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SID\$2/cgcdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SID\$2/cgcdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SID\$2/cgcdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SID\$2/cgcdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SID\$2/cgcdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SID\$2/cgcdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SID\$2/cgcdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SID\$2/cgcdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SID\$2/cgcdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SID\$2/cgcdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SID\$2/cgcdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SID\$2/cgcdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SID\$2/cgcdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SID\$2/cgcdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SID\$2/cgcdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SID\$2/cgcdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C	1	2696.6	82.8	3240	22 AAK52285 Human immune/haema
	2	1568.2	48.1	1666	21 AAC375813 Human ORF1368
	3	1192.8	36.6	7924	24 ABK40069 Human chemically p
	4	1192.8	35.6	7924	24 ABL34138 Human immune syste
C	5	1160.8	35.6	7924	24 ABK40070 Human chemically p
	6	1160.8	35.6	7924	24 ABL34139 Human immune syste
C	7	1160.6	35.6	2021	22 AAS27053 cDNA encoding nove
	8	866.2	26.6	895	20 AAZ24400 Human bladder tumo
	9	534.2	16.4	566	22 AAS27472 cDNA encoding nove
	10	534.2	16.4	566	22 AAS27472 cDNA encoding nove

10	477.4	14.7	481	24	ABK44752
11	438.8	13.5	463	24	ABL66992
12	313.4	9.6	320	16	ABT25457
13	307	9.4	313	24	ABA38341
14	229	7.0	1986	22	ABA09160
15	228.4	7.0	2411	24	ABQ54112
16	217.8	6.7	1161	22	AH755182
17	217.8	6.7	2497	22	AA560878
18	217.8	6.7	2497	22	AA560884
19	217.8	6.7	2497	22	AA560893
20	217.8	6.7	2497	22	AA560895
21	217.8	6.7	2497	23	ABV25781
22	217.8	6.7	2497	23	ABV30037
23	217.8	6.7	3077	22	AHH31912
24	214.2	6.6	866	22	AHH40301
25	202.6	6.2	412	21	AAC01700
26	200.2	6.1	716	21	AAA40104
27	200.2	6.1	861	21	AAA40108
28	198.4	6.1	956	21	AA968887
29	198.4	6.1	1537	21	AAF16196
30	198.4	6.1	3533	22	AHH17889
31	196.8	6.0	674	23	AA571453
32	196.6	6.0	2247	23	ABL29661
33	195.2	6.0	888	21	AAA40109
34	192	5.9	1274	21	AAF18072
35	188.6	5.8	893	23	AA587552
36	185.4	5.7	1540	23	ABL27707
37	174.2	5.3	730	21	AA236696
38	173.4	5.3	1023	21	AA51491
39	173.2	5.3	742	24	ABK83539
40	171.8	5.3	1025	21	AAC34347
41	166.2	5.1	1203	21	AAC34685
42	165.4	5.1	1112	21	AAC46848
43	160.8	4.9	1098	21	AAC48038
44	159.2	4.9	1513	24	ABK71569
45	156.8	4.8	893	23	AA576826

ALIGNMENTS

RESULT 1	
AAK82285/c	
ID	AAK82285 standard; DNA; 3240 BP.
XX	
XX	AAK82285;
AC	
XX	
DT	07-NOV-2001 (first entry)
XX	
DE	Human immune/haematopoietic antigen genomic sequence SEQ ID NO:37097.
XX	
KW	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW	cytostatic; gene therapy; vaccine; metastasis; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO200157182-A2.
XX	
PD	09-AUG-2001.

Pred. No. is the number of results predicted by chance to have a

Db 1180 GGAAGAACCAATCAAAACCTTCTCTCTGAGTTACCAAAACAGGAAACAGCAGGAGG 1121
Qy 2453 GTGCTCAGGACTTAGGACAGGCTATAGCTTAGATGTTGAAAGCAAGAGCAGGAGCA 2512
Db 1120 GTGCTCAGGACTTAGGACAGGCTATAGCTTAGATGTTGAAAGCAAGAGCAGGAGCA 1061
Qy 2513 AGTTGTAATCACTGGCTAATGAGAAAGGAGCAGCTAATCTTAGATGAAGCTGTGAC 2572
Db 1060 AGTTGTAATCACTGGCTAATGAGAAAGGAGCAGCTAATCTTAGATGAAGCTGTGAC 1001
Qy 2573 TAGCTGGAGTGTCTTCTTGAAGATGGGACTCTTGGGTATCAAGACCTATGCCACATC 2632
Db 1000 TAGCTGGAGTGTCTTCTTGAAGATGGGACTCTTGGGTATCAAGACCTATGCCACATC 941
Qy 2633 ACACCTGGGCTAGGAGTGTGATGCTCAGCCCTCAAGTCTGTCTTCAGCAGGACTT 2692
Db 940 ACACCTGGGCTAGGAGTGTGATGCTCAGCCCTCAAGTCTGTCTTCAGCAGGACTT 881
Qy 2693 GAGAAGTTATATTGGGCAAGTGGCTCCAACTCTGTGGCAGTATTTACGTTTCCCTGAAG 2752
Db 880 GAGAAGTTATATTGGGCAAGTGGCTCCAACTCTGTGGCAGTATTTACGTTTCCCTGAAG 821
Qy 2753 ATCAGGAGGCTGCCATTCATCTTCTCTCTAGCCCTCAGGAAGAGACTAT 2812
Db 820 ATCAGGAGGCTGCCATTCATCTTCTCTCTAGCCCTCAGGAAGAGACTAT 761
Qy 2813 ATTTGACTGTACCTAGGGCTTCTGGAAGGAAACATGCAATCAGGATTTCTATAGCT 2872
Db 760 ATTTGACTGTACCTAGGGCTTCTGGAAGGAAACATGCAATCAGGATTTCTATAGCT 701
Qy 2873 GATAGGCCCTATCCACAAGGCCATGACTGGGAAAGATATGGGAGCAGAGGAGATG 2932
Db 700 GATAGGCCCTATCCACAAGGCCATGACTGGGAAAGATATGGGAGCAGAGGAGATG 641
Qy 2933 GATTTTGTAGGTGACGTACCTACCTAACTTTTGTGGCTTGGGCTGCTGCTGAG 2992
Db 640 GATTTTGTAGGTGACGTACCTACCTAACTTTTGTGGCTTGGGCTGCTGCTGAG 581
Qy 2993 GCGCAGAGTTAAGCAGGCTGTGCTGGCTGTGTTACTGCTCAGCAGCTGCTGCTGCT 3052
Db 580 GCGCAGAGTTAAGCAGGCTGTGCTGGCTGTGTTACTGCTCAGCAGCTGCTGCTGCT 521
Qy 3053 GTCCTGAGACTCCATCCAGCCCGCAGCAGCAGCTGCTGCTGAGCTTCTCTCTCC 3112
Db 520 GTCCTGAGACTCCATCCAGCCCGCAGCAGCAGCTGCTGCTGAGCTTCTCTCTCTCC 461
Qy 3113 TGTGACGGGTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3172
Db 460 TGTGACGGGTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 401
Qy 3173 CATCTATTTTAAACACAGATGTTTACAAATAAAGATATTTTCAACACC 3223
Db 400 CATCTATTTTAAACACAGATGTTTACAAATAAAGATATTTTCAACACC 350

RESULT 2

AC75813
ID AAC75813 standard; cDNA; 1666 BP.

XX AC AAC75813;

XX DT 08-FEB-2001 (first entry)

XX DE Human ORFX ORF1368 polynucleotide sequence SEQ ID NO:2735.

XX KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
XX KW vulnerability; antiparkinsonian; nootropic; neuroprotective;
XX KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
XX KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
XX KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
XX KW antiviral; antibacterial; antifungal; antineoplastic; antitumor;
XX KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
XX KW neurodegenerative disorder; osteoarthritis; graft vs host disease;

KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.

XX Homo sapiens.

XX W0200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US08621.

XX 31-MAR-1999; 99US-0127607.

XX 02-APR-1999; 99US-0127636.

XX 05-APR-1999; 99US-0127728.

XX 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

XX WPI; 2000-602362/57.

XX P-PSDB; AAB41604.

XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease -

XX Claim 5; Page 1978-1979; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnerary;
XX antiparkinsonian; nootropic; neuroprotective;
XX osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
XX immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
XX antidiabetic; hypotensive; dermatological; immunosuppressive;
XX antinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
XX antitumor; and antianemic. The sequences can be used for determining
XX the presence of or predisposition to, or preventing or treating
XX pathological conditions associated with an ORFX-associated disorder. The
XX nucleic acids can be used to express ORFX proteins in gene therapy.
XX vectors. The proteins and nucleic acids may be used to treat cancers,
XX proliferative disorders, neurodegenerative disorders, osteoarthritis,
XX graft vs host disease, cardiovascular disease, diabetes mellitus,
XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
XX bacterial or fungal infection, malaria, autoimmune disorders, asthma,
XX allergic, aplastic anaemia, burns, wounds, bone and cartilage damage,
XX nocturnal haemoglobinuria, antinflammatory disease; to enhance
XX coagulation; to inhibit thrombosis; and as a contraceptive.

XX SQ Sequence 1666 BP; 333 A; 529 C; 461 G; 341 T; 2 other;

Query Match 48.1%; Score 1568.2; DB 21; Length 1666;

Best Local Similarity 98.7%; Pred. No. 0;

Matches 1598; Conservative 0; Mismatches 3; Indels 18; Gaps 1;

Qy 2 GCCCGCTGCCCGCCGCGAGTTCCCGGGCCCCCGCTGCCCGAGCAGTACGAGCAGTACGAG 61

Db 48 GCCCGCTGCCCGCCGCGAGTTCCCGGGCCCCCGCTGCCCGAGCAGTACGAGCAGTACGAG 107

Qy 62 TGTGCTGTTCGGCTGTGCTGTGATGCGGGGACTCCGGGGTGGGAGACCTCCCTGCTGTG 121

Db 108 TGTGCTGTTCGGCTGTGCTGTGATGCGGGGACTCCGGGGTGGGAGACCTCCCTGCTGTG 167

Qy 122 CCGCTTCCACGACAGGATTCCTCTCTCCGACATCTCCACCATCGGTGTGACTTTAA 181

Db 168 CCGCTTCCACGACAGGATTCCTCTCTCCGACATCTCCACCATCGGTGTGACTTTAA 227

```
QY 182 GATGAAGACCATAGAGTACAGCGCATCAAAAGTCCGATACAGATCTGGGACACTGCAGG 241
Db 228 GATGAAGACCATAGAGTACAGCGCATCAAAAGTCCGATACAGATCTGGGACACTGCAGG 287
QY 242 GCAGGAGAGATACAGACCATACAAAGCAGTACTATCGCGGGCCCGGAGGATATTTT 301
Db 288 GCAGGAGAGATACAGACCATACAAAGCAGTACTATCGCGGGCCCGGAGGATATTTT 347
QY 302 GGTCTATGATATAGCAGCAGCGCTTTACAGCAGCATCATGAAGTGGGTTCAGTGACGT 361
Db 348 GGTCTATGATATAGCAGCAGCGCTTTACAGCAGCATCATGAAGTGGGTTCAGTGACGT 407
QY 362 GGATGATAGCAGCAGCAGCGCTTCAGAGAGATCTTATTGGGAATAAGCTGATGAGGA 421
Db 408 GGATGATAGCAGCAGCAGCGCTTCAGAGAGATCTTATTGGGAATAAGCTGATGAGGA 467
QY 422 GCAGAAACGGCAGTGGGAAGAGAGCAAGGCGAG-----CAGCTGGC 463
Db 468 GCAGAAACGGCAGTGGGAAGAGAGCAAGGCGAGAAATGTCTTCTTCAGCTGGC 527
QY 464 GAAGGAGTATGGCAGTCTTATGAACAAGTGCCTGCAGCAACCTCAACATTAAGA 523
Db 528 GAAGGAGTATGGCAGTCTTATGAACAAGTGCCTGCAGCAACCTCAACATTAAGA 587
QY 524 GTCAATTCACCGCTGACAGAGCTGGTCTGCAGGCCCATAGNAGGAGCTGGAAGCCT 583
Db 588 GTCAATTCACCGCTGACAGAGCTGGTCTGCAGGCCCATAGNAGGAGCTGGAAGCCT 647
QY 584 CCGGATCGTCCAGCAATAGTGGCACTGGCAGAGCTGGAGGAGGAGGAGCAAC 643
Db 648 CCGGATCGTCCAGCAATAGTGGCACTGGCAGAGCTGGAGGAGGAGGAGGAGCAAC 707
QY 644 CGAGGCGCCAGCACTCTTCGAAACCTGCTGGTGTGAGTCTGTGTGGGGCACCCCA 703
Db 708 CGAGGCGCCAGCACTCTTCGAAACCTGCTGGTGTGAGTCTGTGTGGGGCACCCCA 767
QY 704 CAGACACACCTCTTCCCTCAGAGGCGCTGGGACAGAGGAGGAGCGGGCTTGGCC 763
Db 768 CAGACACACCTCTTCCCTCAGAGGCGCTGGGACAGAGGAGGAGCGGGCTTGGCC 827
QY 764 TGCTGTGTCTCTCTGTGTGATGACCCCTATTGAGTATGATAGCCACTACTCCCGCTGCC 823
Db 828 TGCTGTGTCTCTCTGTGTGATGACCCCTATTGAGTATGATAGCCACTACTCCCGCTGCC 887
QY 824 TGGCCCTGAGAGGGCTCTCTGTATCTCAAGAGCCCTGTGCCAGCCGCTGCACCC 883
Db 888 TGGCCCTGAGAGGGCTCTCTGTATCTCAAGAGCCCTGTGCCAGCCGCTGCACCC 947
QY 884 TGGAGTGTCTCTTTCAGCCTGTTTCCCGCAGCCAGGCGCTGTACGACCCCGCAGTGT 943
Db 948 TGGAGTGTCTCTTTCAGCCTGTTTCCCGCAGCCAGGCGCTGTACGACCCCGCAGTGT 1007
QY 944 GCGCAAGCACTGTCTCACCATCCCGCAGCCAGCAGCAACAGCAGCGGCTGGAGTCCAG 1003
Db 1008 GCGCAAGCACTGTCTCACCATCCCGCAGCCAGCAGCAACAGCAGCGGCTGGAGTCCAG 1067
QY 1004 GGCATTTAGTGTCTCTTTCGCGGATCGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 1063
Db 1068 GGCATTTAGTGTCTCTTTCGCGGATCGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 1127
QY 1064 CCCCACTCTCTCTCTGACCCCTCCCTCCCGTGGTTCGTATCAAGAGCTCTCAAA 1123
Db 1128 CCCCACTCTCTCTCTGACCCCTCCCTCCCGTGGTTCGTATCAAGAGCTCTCTCAAA 1187
QY 1124 CCCCGTCCCGGTGTCTCTGTGTGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1183
Db 1188 CCCCGTCCCGGTGTCTCTGTGTGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1247
QY 1184 CAAGGGATGAGCCAGGCTGTGGGAGGTTCCACCCCTTGGATCCAGGAAGAACCTCC 1243
Db 1248 CAAGGGATGAGCCAGGCTGTGGGAGGTTCCACCCCTTGGATCCAGGAAGAACCTCC 1307
QY 1244 ACCCTGCTGTGGGTGGGCCAAGGCTACAGGCTGTCTCTCTCTCTCTCTCTCTCTCTCT 1303
```

RESULT 3

ABK40069

ID ABK40069 standard; DNA; 7924 BP.

XX AC ABK40069;

XX XX

DT 21-MAY-2002 (first entry)

XX Human chemically pretreated gene sequence #76 strand 1.

Human; ds; bisulphite treatment; CpG; DNA methylation; cancer; tumour;
cytostatic; ALDH6; CYP11A; CYP11B1; CYP3A3; DPYD; EPHX2; OCLN; TXNRD1;
UGT8; MRP; pharmacogenomics; SNP; single nucleotide polymorphism.

XX Homo sapiens.

XX WO200202806-A2.

XX 10-JAN-2002.

XX 29-JUN-2001; 2001WO-EP07470.

XX 30-JUN-2000; 2000DE-1032529.

XX 01-SEP-2000; 2000DE-1043826.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2002-154757/20.

XX New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,
useful for detecting cytosine methylation state of genes associated
with pharmacogenomics and for therapy of diseases e.g. cancer

XX Claim 1; SEQ ID No 151; 24pp; English.

XX The invention relates to a nucleic acid comprising a sequence at
least 18 bases in length of a segment of the chemically pretreated DNA
of genes associated with pharmacogenomics according to one of the
sequences of the genes ALDH6 (NM_00693), CYP11A (NM_000781), CYP11B1
(NM_000497), CYP3A3 (NM_000776 and NM_017450), DPYD (NM_000110), EPHX2
(NM_001979), OCLN (NM_002538), TXNRD1 (NM_003330), UGT8 (NM_003360),
MRP (NM_004996, NM_019900, NM_019901, NM_019902, NM_019862, NM_019898,
NM_019899) and their complementary sequences, or a sequence (S1) chosen
from 87 sequences and their complements. The chemical pretreatment
is bisulphite treatment to convert cytosines (but not methyl-cytosines)
into uracils. Also included are an oligomer (II) in particular an
oligonucleotide or a peptide nucleic acid (PNA)-oligomer, comprising in

Db 1861 TGTGTTTCGGGTTATATATGAATGTGAGTAGGGTTATTTATTTAAATATAGATGTT 1920

QY 3196 TACAAAATAAGATTATTTCAACACCAC 3223

Db 1921 TATATAAATAAGATTATTTAAATATTC 1948

RESULT 4

ABL34138

ID ABL34138 standard; DNA: 7924 BP.

AC ABL34138;

XX

XX 26-MAR-2002 (first entry)

DE Human immune system associated gene SEQ ID NO: 2111.

XX Human; immune system disease; cytosine methylation; antiasthmatic;

KW antiarteriosclerotic; antiandemic; cytosine; nontropic;

KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;

KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;

KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;

KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;

KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;

XX gene; ds.

XX Homo sapiens.

OS

XX W0200200928-A2.

PN

XX 03-JAN-2002.

PD

XX 02-JUL-2001; 2001WO-EP07537.

PF

XX 30-JUN-2000; 2000DE-1032529.

PR

XX 01-SEP-2000; 2000DE-1043826.

PA (EPIC-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

PI WPI; 2002-130909/17.

DR

XX Nucleic acid comprising fragment of chemically modified gene, useful

PT for diagnosis and treatment of diseases associated with abnormal

PT cytosine methylation -

XX Claim 1; SEQ ID NO 2111; 32pp + Sequence Listing; German.

PS

XX The present invention provides a number of human immune system associated

CC genes which are modified by the methylation of cytosines. The sequences

CC can be used in the diagnosis and treatment of immune system disorders,

CC including eye diseases such as retinopathy, neovascular glaucoma and

CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid

CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,

CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel

CC diseases. The present sequence is a gene of the invention.

XX

SQ Sequence 7924 BP; 2087 A; 83 C; 2037 G; 3717 T; 0 other;

Query Match 36.68; Score 1192.8; DB 24; Length 7924;

Best Local Similarity 75.8%; Pred. No. 5e-279;

Matches 1476; Conservative 0; Mismatches 472; Indels 0; Gaps 0;

QY 1276 GGTGCTTCTCTCCACCCACCCACCTGCTCCTCATGTGCGATGGCGCTGCTCC 1335

Db 1 GGTGTTTTTTTTTTTTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 60

QY 1336 AGTGACCTGGAAAGTGGAGCATCGAGGTAGGAGGAAACAGCAACCGGGGAGTCTCGA 1395

Db 61 AGTGATTGGGAAGTGGAGTATCGAGTAGGAGGGAACGGTAATTAGGAGCTTTTCA 120

QY 1396 GCCTGGGGCTGCCTACCTCTACCCATTCGCCGACAGAGCTTTGCCCTTGCCTGCTG 1455

Db 121 GTTTGGGGTGTGTTTATTTATTTATTTTCGATTAGATTGTTGTTGTTGTTGTTGTT 180

QY 1456 CCGCCTGCCCTCTTTTGGGAACTGAGCTGCAGACGAGCTGCTCAGAGAAGGAACAAAT 1515

Db 181 TCGTTTGTGTTTTTTGGGAAATGAGTTTAGAGGTAGGTGTTTAGAGAAGGAATAAAT 240

QY 1516 GAGGGTGGCAGGATAAAAAGTACCTCCATTCTCTACCTCCCATCCAGCATGAACACA 1575

Db 241 GAGGGTGGTAGGATAAAAAGTTATTTTATTTTATTTTATTTATCTAGTATGAATATA 300

QY 1576 ATTCTCTCCACCTGGCTCCCAATTTAAAGATGTGGACCAAGCCCTGTGGGTACTCCAG 1635

Db 301 ATTTTATTTTATTTGTTTAAATTTAAAGATGTGGATTAAAGTTTGTGGTATTTTAG 360

QY 1636 GGGCAAGGAGAGCCCTGGGCTCAGTGACACTGTCAGGCCCAACCATCACCTCCCAAGGG 1695

Db 361 GGTAGGAGAGTTTGGGGTTAGTGATATTTAGTTTAAATATTTATTTATTTAAAGGG 420

QY 1696 GAGCATTTGGAATGAAGGACTAGCTCTATGTATCATCAGTTAAAGCAAGGGAGAGCTGG 1755

Db 421 GAGTATTTGGAATGAAGGATTAGTTTATTTATTTAGTTAAAGTTAAAGGAGAGTTGG 480

QY 1756 CAGGGACACAGTTTCCACAGCAGAGGGGAATGTAGCACAGCAGAGGCCCTCTAGGCC 1815

Db 481 TTAGGGATAGTAGTTTGTATAGTAGAGGGAATGTAGTAATAGTAGGTTTATTTAGTTT 540

QY 1816 CATCTTCCATTTCTTAGTAAAGAGCATTTCTCAGACTCCCGAGCGGAGGACTCAGC 1875

Db 541 TATTTTATTTTATTTAGTAAAGAGTATTTTATTTAGATTTTATTTAGCGGAGGATGAGT 600

QY 1876 CTAGCCTTCAGCAACCAAGTTTCTCTGGGACCCAAAGTTTATGGGAGAAGGGCAAGAC 1935

Db 601 TTAGTTTATTTAGTAATTAAGCTTTTGGGATTAAAGTTTATGGGAGAAGGGTAAAGAT 660

QY 1936 TTCAATGGGAAGAGAGAGGAGCCCTGGGTAGAAACCTTGGTCTCTCTTTGGCC 1995

Db 661 TTTATGGGAAGAGAGAGAGGTTTGGGTAGAAACCTTGGTCTCTCTTTTGGTT 720

QY 1996 TTTAAGACAAAGGCTCATCTTGCCCTCTACCTCTCATAGGCTTGGGTTTCCCAACC 2055

Db 721 TTTAAGATAAAGCGTTTATTTGTTTTTATTTTATTTTATTTATTTTATTTTATTT 780

QY 2056 ACATGTGCTACAGTGGAGGAGAGGACTCCTCTCCAGAGTCTATGTTTCAGGAA 2115

Db 781 ATATTGGTTATAGGTGGAGGAGAGGATTTTATTTTATTTAGAGTCTTATGTTTAGGAA 840

QY 2116 GTTCTTTAACCCCAATATGCCCCAAGAGTAGTCTGTAGGAGGCCCTTTAAGACGGAACA 2175

Db 841 GTTTTATTTTATTTATTTATTTTAAAGATAGTTCTGTAGGAGTTTAAAGACGGAATA 900

QY 2176 AGTAATTTACAGTTCTACTGGGTTCTCTGCCACCTGCCCAAGTGGCGGAGCCCTAGG 2235

Db 901 AGTAATTTATTTATTTTATTTGTTTATTTTATTTATTTTATTTTATTTTATTTTATTT 960

QY 2236 AAGAGGTCTATTTTAAAGCAGCATTTAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2295

Db 961 AAGAGGTCTATTTTAAAGTTATATATAGTTTGTATTGCTGTTGTTGTTTAAATAAAGA 1020

QY 2296 ACTGGGTGTTGAGTATTTCACTAAGAACCAAAATCCAGGCGCACTCATATGTGAAGGA 2355

Db 1021 ATTGGGTGTTGAGTATTTTATTTAATTAAGAATTTAAATTTAGGGTATTTATATGTGAAGGA 1080

QY 2356 TAAGAACCTCACTCTCTTCTCTCTCCCAAAAGAGTGGGGAAGCAACCATCAACCTTTC 2415

Db 1081 TAAGAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1140

QY 2416 CTCTGACTTACCAACACAGGAAACAGCAGAGAGGTTGCTCAGGACTTTAGGACAGG 2475

Db 1141 TTTTGTATTTTATTAATTAAGAAATAGTAGGAGGTTGCTTTAGGATTTAGGATAGG 1200

QY 2476 GTATAGCTTATAGTGTGGAAGCAAGAGAGAGAGGAGTTGTAATACCTGCTGCTAATGA 2535

Db 7562 AAAAAAAAAAACCCTAAATCAATACACATATCAACCAACCATACACTCCACAAAAA 7503
Qy 1698 GCATTGGGAATGAAGACTAGCTCCTATGATATCAGGTTAAGAGCAAGGAGCTGGCC 1757
Db 7502 ACATTAAAAATAAAAAATCTCTATATATCAAAATTAACAAAAAATAAACCTAAC 7443
Qy 1758 AGGACAGCAGTTTGCACACAGCAGGGGATGTAGCAACAGCAGGCGCTCCTAGGCCCA 1817
Db 7442 AAAAAACAACATTTACACAACAACAAAAAATATAACACAACAACAAAAACCTCTAAACCCCA 7383
Qy 1818 TCTTCCATTTCTTAGGTAAAGAGAGCATTTCTTCAGACTCCAGGCGAGGAGCTGAGCCT 1877
Db 7382 TCTTCCATTTCTTAATAAATAAACAATTTCTTCACACTCCCAACGAAAACTAACCT 7323
Qy 1878 AGCCTTCAGCAACAAGTTTCTCTGGGACCCAAAGTTTATGGGAGAGGCGCAAGACTT 1937
Db 7322 AACCTTCAACAACCAAAATTTCTCTAAACCCCAAAATTTATAAAAAAATAACAAAACTT 7263
Qy 1938 CATGGGAGAGAGAGGAGGCGCTGGGTAGAACGCTTGGTCTGTCTTCTTTGGCCTT 1997
Db 7262 CATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7203
Qy 1998 TAAGACAAGCGCTCATTTGGCCTCTACCTCTGATAGGCTTGGGTTTGGCAACAC 2057
Db 7202 TAAACAACAAGCTCATTTACCTCTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7143
Qy 2058 ACTGTGGCTACAGGTGGAGGAGAGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2117
Db 7142 ACTATAACTACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7083
Qy 2118 TTCTTTAACCCCATATGGCCAGAGTAGTCTCTAGGAGGCTTTTAAAGACGAGCAAG 2177
Db 7082 TTCTTTAACCCCATATGAACCAAAATAAATCTCTGTAATAAATCTCTCTCTCTCTCTCT 7023
Qy 2178 TAATTTACCAAGTTCTACTGGGTTCTCTGCCACCGTCCCAAGTGGCGGAGGCTTAGGAA 2237
Db 7022 TAATTTACCAATTTCTACTAAATTTCTACCCCGTCCCAATTAACGAACTTAAAA 6963
Qy 2238 GAGGTCATTTCTTAAGCCACACATATAGCTGACATGGGTGGCTGCAGCCCAACAAAGAC 2297
Db 6962 AAAAAATCATTTTAAACACACATTAACCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAG 6903
Qy 2298 TGGGTGTGAGTATTCATCACTAAGACCAAAATCCAGGCACTCATATGTGAGGATA 2357
Db 6902 TAAATATTAATATTCATCACTAAGACCAAAATCCAAACACTCATATATAAATAA 6843
Qy 2358 AGAAGCTCACTTCTTACTCTCCAAAAAGAGTGGGGAAGAACATCAACCTTTCTCT 2417
Db 6842 AAAAACTCACTTCTTACTCTCCAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 6783
Qy 2418 CTTGACTTACCAACCAAGGAAAAACAGCAGGAGGGTGGCTCAGGACTTAGGACAGGGT 2477
Db 6782 CTTAACTTACCAACCAAGGAAAAACAAAAAATAAATAAATAAATAAATAAATAAATAA 6723
Qy 2478 ATAGCTTAGATGTGGAAACCAAGGAGCAGGAGTTGTAATCACTGCTTAAGAGA 2537
Db 6722 ATAACTTAAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6663
Qy 2538 AAGGAGACAGCTAACTTCTAGGATGAAGCTGTGACTAGGCTGGGTTGCTTCTTTGAAGA 2597
Db 6662 AAAAAACAACCTTAACTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6603
Qy 2598 TGGGACTTCTTGGGTATCAAGACTATGCCACATCACTAGGCGCTAGGGAAGTAGTGA 2657
Db 6602 TAAACCTCTTAAATATCAAAACCTATACCACATCACACTTAAACCTTAAATAAATAA 6543
Qy 2658 TGGCAGCCCTCAAGTCTGTCTTCCAGCCAGGAGCTTGAGAGTTATATTGGGAGTGGCTC 2717
Db 6542 TACCAACCCCTCAATCTCTCTCAACCAAAAACTTAAAAAATAAATAAATAAATAAATAA 6483
Qy 2718 CAATCTGTGACAGTATTTCACTTTCCCTGAAGATCAGGCGGTTGCCATTCATTTGTC 2777
Db 6482 CAATCTAATAACCAATATTTCACTTTCCCTTAAATCAAAATAAATAAATAAATAAATAA 6423

Qy 2778 TTCTCTCTCTAGCCCTCAGAAAGAGGACTATATTTGTACTGTACCTTAGGGTTCT 2837
Db 6422 TTCTCTCTCTAACCCTCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 6363
Qy 2838 GGAAGGAAAAACATGAGTATAGACTATAGACTATAGACTATAGACTATAGACTATAG 2897
Db 6362 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6303
Qy 2898 GACTGGGAAAGGTATGGGAGCAGAGGAGATTTGGATTTTAGGGTGACGTACGTCA 2957
Db 6302 AACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6243
Qy 2958 CCTAAACTTTTGGTGGCCTGGGCGCATGTCTTGAAGCCAGACTGTTTAAAGCAGCTCTGC 3017
Db 6242 CCTAAACTTTTAAATAACCTTAAACATATCTTAAACCAAACTATTTAAACCAAACTCTAC 6183
Qy 3018 TGCCCTGTTTACTCGTCGACCACTCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3077
Db 6182 TAACCTTATTTACTCGTCGACCACTCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6123
Qy 3078 GCAGCCACCTGCTGAGCCTCCACTATCTCCTGCTGAGCGGTGAATTCGTTGACTG 3137
Db 6122 ACAGCCACCTGCTGAGCCTCCACTATCTCCTATAGGATTAACCTTCTGATACTA 6063
Qy 3138 TGCTCGGGTCCATATATGAATTTGTGAGCGGTTCATCTATTTTAAACACAGATGTTA 3197
Db 6062 TATCTCGAATCCATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6003
Qy 3198 CAAAAATAAGATTTATTCACACCAACCAAAAAAATAAATAAATAAATAAATAAATAA 3257
Db 6002 CAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5943

RESULT 6

ABL34139/C
ID ABL34139 standard; DNA; 7924 BP.

XX ABL34139;

DT 26-MAR-2002 (first entry)

DE Human immune system associated gene SEQ ID NO: 2112.

DE Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.

XX Homo sapiens.

XX WO200200928-A2.

XX 03-JAN-2002.

XX 02-JUL-2001; 2001WO-EP07537.

XX 30-JUN-2000; 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1043826.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2002-130909/17.

XX Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -

17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-465460/50.
 DR P-PSDB; NA017136.
 XX
 Novel polypeptides useful for diagnosing, treating, preventing and/or
 PT prognosing disorders related to the proteins, including cancers, immune
 PT disorders and neuronal disorders
 XX
 PS Claim 1; SEQ ID No 88; 880pp; English.
 XX
 The invention relates to novel isolated polypeptides (I), and
 CC polynucleotides (II). (i), (ii) and the antibody to (i) are useful for
 CC diagnosing, preventing and treating diseases including immune system
 CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
 CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
 CC transplant rejections and graft versus host disease, infectious diseases
 CC other blood-related disorders (sickle cell anaemia), myeloproliferative
 CC disorders, primary haematopoietic disorders, hyperproliferative
 CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
 CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
 CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
 CC disorders (e.g. glomerulonephritis), cardiovascular disorders
 CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
 CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
 CC Addison's disease), reproductive system disorders, gastrointestinal
 CC disorder (inflammatory disorders), liver disorders (cirrhosis),
 CC as stimulators of B-cell responsiveness to pathogens, activators of
 CC T-cells, to induce higher affinity antibodies, and as a means to induce
 CC tumour proliferation in pathologies e.g. acquired immune deficiency
 CC syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction
 CC pathway protein coding sequences and PCR primers of the invention.
 XX
 Query Match 35.6%; Score 1160.6; DB 22; Length 2021;
 Best Local Similarity 99.2%; Pred. No. 1.8e-271;
 Matches 1166; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 QY 2 GCCCGCTGCCGCGCAGTTCGCCGCCCGCTGCCGCCAGTCATGGCGGAGCAGTACGA 61
 DB 38 GCCCGCTGCCGCGCAGTTCGCCGCCCGCTGCCGCCAGTCATGGCGGAGCAGTACGA 97
 QY 62 TGTGCTGTTCCGCGCTGCTGCTGATCGGGGACCTCCGGGGTGGGCAAGACCTGCTGTG 121
 DB 98 TGTGCTGTTCCGCGCTGCTGCTGATCGGGGACCTCCGGGGTGGGCAAGACCTGCTGTG 157

122 CCGCTTACCGACAACGAGTTCCACTCCTCGCACATCTCCACCATCGGTGTGACTTTAA 181
 DB 158 CCGCTTACCGACAACGAGTTCCACTCCTCGCACATCTCCACCATCGGTGTGACTTTAA 217
 QY 182 GATGAAGACCATAGAGGTAGAGCGCATCAAGTGGGATACAGATCTGGGACACTGCAGG 241
 DB 218 GATGAAGACCATAGAGGTAGAGCGCATCAAGTGGGATACAGATCTGGGACACTGCAGG 277
 QY 242 GCAGGAGATACCGACCATCAAAAGCAGTACTATCGCGGGCCCGAGGGATATTTT 301
 DB 278 GCAGGAGATACCGACCATCAAAAGCAGTACTATCGCGGGCCCGAGGGATATTTT 337
 QY 302 GGTCTATGACATTAGCAGCGAGCGCTTTACAGCAGATCATGAAGTGGTCAAGTCAAGT 361
 DB 338 GGTCTATGACATTAGCAGCGAGCGCTTTACAGCAGATCATGAAGTGGTCAAGTCAAGT 397
 QY 362 GGATGAGTACCGACCATCAAAAGCAGTACTATCGCGGGCCCGAGGGATATTTT 421
 DB 398 GGATGAGTACCGACCATCAAAAGCAGTACTATCGCGGGCCCGAGGGATATTTT 457
 QY 422 GCAGAAACGGCAGGTGGGAAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 481
 DB 458 GCAGAAACGGCAGGTGGGAAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 517
 QY 482 CTTCTATGAACAAAGTGCCTGACCAACCTCAACATTAAAGAGTCAATTACGGGTCTGAC 541
 DB 518 CTTCTATGAACAAAGTGCCTGACCAACCTCAACATTAAAGAGTCAATTACGGGTCTGAC 577
 QY 542 AGAGCTGTGCTGCAGGCCCATAGGAAGGAGTGGGAAGGCTCGCGATGCGTGCAGCAA 601
 DB 578 AGAGCTGTGCTGCAGGCCCATAGGAAGGAGTGGGAAGGCTCGCGATGCGTGCAGCAA 637
 QY 602 TGAGTTGGCACTGGCAGAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 661
 DB 638 TGAGTTGGCACTGGCAGAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 697
 QY 662 TTCGAAACCTGCTGGTGTGCTGAGTCTGTGTGGGGGACCCACACAGGACCCCTCTTCCC 721
 DB 698 TTCGAAACCTGCTGGTGTGCTGAGTCTGTGTGGGGGACCCACACAGGACCCCTCTTCCC 757
 QY 722 TCAGGAGGCCGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 781
 DB 758 TCAGGAGGCCGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 817
 QY 782 TGATGAACCTATGATGATATAGTACAGTACTACTCCCTGCTGCTGCTGCTGCTGCTGCTGCT 841
 DB 818 TGATGAACCTATGATGATATAGTACAGTACTACTCCCTGCTGCTGCTGCTGCTGCTGCTGCT 877
 QY 842 TGCTGTCTATCTCAAGCAGCCCTGTCGCCAGCCGCTCCACCTGAGTGGTCTTCTTCAG 901
 DB 878 TGCTGTCTATCTCAAGCAGCCCTGTCGCCAGCCGCTCCACCTGAGTGGGCTTCTTCAG 937
 QY 902 CTTGTTTCCCGACGACAGGCTGTCTAGCAGCCCGGATGTGCGGCAAGCACTGTCTCA 961
 DB 938 CTTGTTTCCCGACGACAGGCTGTCTAGCAGCCCGGATGTGCGGCAAGCACTGTCTCA 997
 QY 962 CCATCCCGACCCAGCAGACAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1021
 DB 998 CCATCCCGACCCAGCAGACAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1057
 QY 1022 TTTCTCCGTCATCGTGTCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1081
 DB 1058 TTTCTCCGTCATCGTGTCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1117
 QY 1082 GACCCCTCCCTCCGCGTGGTTCGTATCAAGTCTCTCAAAAGTCTCTCAAAAGTCTCTCTCTCT 1141
 DB 1118 GACCCCTCCCTCCGCGTGGTTCGTATCAAAAGTCTCTCAAAAGTCTCTCTCTCTCTCTCTCT 1177
 QY 1142 CTGCTGTGTCAGTCTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1176
 DB 1178 CTGCTGTGTCAGTCTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1212

PD 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US01312.
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226581.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0233397.
PR 14-SEP-2000; 2000US-0233398.
PR 14-SEP-2000; 2000US-0233399.
PR 14-SEP-2000; 2000US-0233400.
PR 14-SEP-2000; 2000US-0233401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.

PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2000US-0254097.
XX 2001US-0259678.
PA (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-465460/50.
XX

DR P-PSDB: AAU17555.

XX Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders

XX Claim 1; SEQ ID No 507; 880pp; English.

XX The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorders), liver disorders (cirrhosis),
CC as stimulators of B-cell responsiveness to pathogens, activators of
CC T-cells, to induce higher affinity antibodies, and as a means to induce
CC tumour proliferation in pathologies e.g. acquired immune deficiency
CC syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction
CC pathway protein coding sequences and PCR primers of the invention.

Query Match 16.4%; Score 534.2; DB 22; Length 566;
Best Local Similarity 98.9%; Pred. No. 1.3e-119;
Matches 533; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 GCCCGCTGCCGCGCCGAGTTCGCGCCCGCTGGCCCGCCAGTATGCGGAGCAGTACGA 61
Db 28 GCCCGCTGCCGCGCCGAGTTCGCGCCCGCTGGCCCGCCAGTATGCGGAGCAGTACGA 87
QY 62 TGTGCTGTTCGGCTGCTGCTGATCGGGGACTCGGGGCTGGGCAAGACCTGCTGCTG 121
Db 88 TGTGCTGTTCGGCTGCTGCTGATCGGGGACTCGGGGCTGGGCAAGACCTGCTGCTG 147
QY 122 CCGCTTCCAGCAGCAACGAGTTCCACTCCTCGCACATCTCCACCATCGCTGCTGCTAA 181
Db 148 CCGCTTCCAGCAGCAACGAGTTCCACTCCTCGCACATCTCCACCATCGCTGCTGCTAA 207
QY 182 GATGAAGACCATAGAGTAGACGCGCATCAAGTCCGATACAGATCTGGCAGCTGCGAG 241
Db 208 GATGAAGACCATAGAGTAGACGCGCATCAAGTCCGATACAGATCTGGCAGCTGCGAG 267
QY 242 GCAGGAGATACACAGACCATCAAGAGCAGTACTATCGCGGGCCCGAGGGATATTTT 301
Db 268 GCAGGAGATACACAGACCATCAAGAGCAGTACTATCGCGGGCCCGAGGGATATTTT 327
QY 302 GGTCTATGACATTAGCAGCGGCTCTTACCAGCATCATGAAGTGGGTGACGTACGT 361
Db 328 GGTCTATGACATTAGCAGCGGCTCTTACCAGCATCATGAAGTGGGTGACGTACGT 387
QY 362 GGATGAGTAGCAGCAGGAGCGCTCCAGAGATCCCTATTGGGAATAGGCTGATGAGGA 421
Db 388 GGATGAGTAGCAGCAGGAGCGCTCCAGAGATCCCTATTGGGAATAGGCTGATGAGGA 447
QY 422 GCAGAAACGCGAGTGGGAAGAGAGCAAGGCGAGCAGCTGGCGAAGGAGTATGCGATGGA 481
Db 448 GCAGAAACGCGAGTGGGAAGAGAGCAAGGCGAGCAGCTGGCGAAGGAGTATGCGATGGA 507
QY 482 CTTCTATGAACAGTGCCTTGCACCAACCTCAACATTAAGAGTCAFTTCACGCGTCTGA 540
Db 508 CTTCTATGAACAGTGCCTTGCACCAACCTCAACATTAAGAGTCAFTTCACGCGTCTGA 566

RESULT 10

ABK44752
XX ID ABK44752 standard; cDNA; 481 BP.
XX AC ABK44752;
XX
XX 05-JUN-2002 (first entry)
XX
XX cDNA encoding colon tumour protein, SEQ ID No 303.
XX Human; colon tumour; vaccine; colon cancer; immunogenic;
XX immunotherapy; gene; ss.
XX Homo sapiens.
XX WO200212328-A2.
XX
XX 14-FEB-2002.
XX
XX 31-JUL-2001; 2001WO-US24218.
XX
XX 03-AUG-2000; 2000US-223283P.
XX 28-MAR-2001; 2001US-279763P.
XX 29-JUN-2001; 2001US-302051P.
XX (CORI-) CORIXA CORP.
XX King GE, Meagher MJ, Xu J, Secretist H;
XX WPI; 2002-241739/29.
XX
XX New colon cancer polypeptides and polynucleotides, useful as vaccines,
XX for diagnosing, preventing, and treating colon cancer, and as markers
XX for the progression of cancer
XX
XX Claim 1; SEQ ID No 303; 147pp; English.
XX
XX The invention relates to polynucleotides encoding colon tumour proteins.
XX The polynucleotides and encoded polypeptides are useful in pharmaceutical
XX compositions, such as vaccines, for the diagnosis, prevention, and
XX treatment of colon cancer. Polynucleotide sequences may be used as
XX hybridisation probes or primers, and in the design and preparation of
XX ribozyme molecules for inhibiting expression of tumour polypeptides and
XX immune response against cancer. The compositions are useful for stimulating an
XX colon cancer, and as markers for the progression of cancer.
XX ABK4450-ABK46237 represent coding sequences of human colon tumour
XX proteins of the invention.
XX Note: With the exception of SEQ ID No 1 and 2, the sequence data
XX for this patent did not form part of the printed specification but was
XX supplied by the European Patent Office.

XX Sequence 481 BP; 111 A; 113 C; 130 G; 125 T; 2 other;

Query Match 14.7%; Score 477.4; DB 24; Length 481;
Best Local Similarity 99.4%; Pred. No. 7.4e-106;
Matches 478; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2582 GTTGCTTCCTTGAAGATGGGACTCTTGGGTATCAAGCTATCCACCATCAGTGGG 2641
Db 1 GTTGCTTCCTTGAAGATGGGACTCTTGGGTATCAAGCTATCCACCATCAGTGGG 60
QY 2642 CTAGGGAAGTAGGTGATGCCAGCCCTCAAGTCTCTTTCAGCCAGGACTTGAGAAGTTA 2701
Db 61 CTAGGGAAGTAGGTGATGCCAGCCCTCAAGTCTCTTTCAGCCAGGACTTGAGAAGTTA 120
QY 2702 TATTGGCAGTGGCTCCCAATCTGTGGACCAAGTATTCAGCTTCCCTTGAAGATCAGG 2761
Db 121 TATTGGCAGTGGCTCCCAATCTGTGGACCAAGTATTCAGCTTCCCTTGAAGATCAGG 180
QY 2762 GGTGCCATTATCTCTTCTCTCTAGCCCTCAGGAAGAGGACTATATTGTTACT 2821
Db 181 GGTGCCATTATCTCTTCTCTCTAGCCCTCAGGAAGAGGACTATATTGTTACT 240

QY 2822 GTACCTAGGGTTCTGGAAGGAAAAACATGGAATCAGGATTCTATAGACTCATAGGCC 2881
 |||||
 Db 241 GTACCTAGGGTTCTGGAAGGAAAAACATGGAATCAGGATTCTATAGACTCATAGGCC 300
 |||||
 QY 2882 TATCCACAGGCCATGACTGGGAAAGTATGGAGCAGAGGAGAAATGGGATTTAG 2941
 |||||
 Db 301 TATCCACAGGCCATGACTGGGAAAGTATGGAGCAGAGGAGAAATGGGATTTAG 360
 |||||
 QY 2942 GGTGAGCTACCTCACCTAACTTTGGTGGCTGGGCAATGCTTTGAGGCCCAGACT 3001
 |||||
 Db 361 GGTGAGCTACCTCACCTAACTTTGGTGGCTGGGCAATGCTTTGAGGCCCANACT 420
 |||||
 QY 3002 GTTAAGCAGGCTCTGCTGGCCCTGTTTACTGCTACCACTCTGACCTGCTGTTGAGA 3061
 |||||
 Db 421 GTTAAGCAGGCTCTGCTGGCCCTGTTTACTGCTACCACTCTGACCTGCTGTTGAGA 480
 |||||
 QY 3062 C 3062
 Db 481 C 481

RESULT 11
 ABL66992/C
 ID ABL66992 standard; DNA; 463 BP.
 AC ABL66992;
 XX
 XX 15-MAY-2002 (first entry)
 XX
 XX Thyroid cancer related gene sequence SEQ ID NO:5329.
 XX
 XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.
 XX
 XX Homo sapiens.
 XX
 XX WC200194629-A2.
 PN
 PD 13-DEC-2001.
 XX
 XX 30-MAY-2001; 2001WO-US10838.
 XX
 XX 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 20-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.

PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237423P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 XX
 XX (AVAL-) AVALON PHARM.
 XX
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX
 DR WPI; 2002-188264/24.
 XX
 XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 XX
 PS Claim 1; SEQ ID 5329; 44pp; English.
 XX
 CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 847 sequences (given in ABL61864
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilm's tumour.
 XX
 SQ Sequence 463 BP; 116 A; 108 C; 110 G; 129 T; 0 other;
 Query Match 13.5%; Score 438.8; DB 24; Length 463;
 Best Local Similarity 99.3%; Pred. No. 1.7e-96;
 Matches 451; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
 QY 2790 CCCCCTCAGGAAGAAGGACTATATTTGCTACCTACCTAGGGTTCTGGAGGAAAAC 2849
 |||||
 Db 453 CCCCCTCAGGAAGAAGGACTATATTTGCTACCTACCTAGGGTTCTGGAGGAAAAC 394
 |||||
 QY 2850 ATGGAATCAGGATTCTATAGACTGATAGCCCTATCCACAGGGCCATGACTGGGAAAAG 2909
 |||||
 Db 393 ATGGAATCAGGATTCTATAGACTGATAGCCCTATCCACAGGGCCATGACTGGGAAAAG 334
 |||||
 QY 2910 GTATGGGAGCAGAGGAGAAATTTGGGATTTTAGGTGTCAGCTACGCTACCCCTAACTTTT 2969
 |||||
 Db 333 GTATGGGAGCAGAGGAGAAATTTGGGATTTTAGGTGTCAG-TACGCTACCCCTAACTTTT 275
 |||||
 QY 2970 GGTGGCCTGGGCGATGCTTTGAGGCCCGACACTGTTAAGCAGGCTCTGCTGGCCTTTAC 3029
 |||||
 Db 274 GGTGGCCTGGGCGATGCTTTGAGGCCCGACACTGTTAAGCAGGCTCTGCTGGCCTTTAC 215
 |||||
 QY 3030 TCGTCACCACTCTGCACCTGCTGCTTTGAGACTCCATCCAGCCCCCAGCAGCCACCTG 3089
 |||||
 Db 214 TCGTCACCACTCTGCACCTGCTGCTTTGAGACTCCATCCAGCCCCCAGCAGCCACCTG 155
 |||||
 QY 3090 CTCCTGAGCCCTCCACTATCTCCCTGTGACGGGTGAACCTTCGTGTACTGTGTCTCGGGTCC 3149
 |||||

```
Db 154 CTCCTGAGCCTCCACTATCTCCCTGTGACGGGTGAACCTTGGTACTGTCTCGGGTCC 95
Qy 3150 ATATATGAATGTGAGCAGGGTTCATCTATTTTAAACACAGATGTTTACAAAAATAAGAT 3209
Db 94 ATATATGAATGTGAGCAGGGTTCATCTATTTTAAACACAGATGTTTACAAAAATAAGAT 35
Qy 3210 TATTTCACCAACCCACAAAAAATAAAAAAATAAAAAA 3243
Db 34 TATTTCACCAACCCACAAAAAATAAAAAAATAAAAAA 1

RESULT 12
AAT25457
ID AAT25457 standard; cDNA to mRNA; 320 BP.
XX AC
XX AAT25457;
XX DT- 17-OCT-1996 (first entry)
XX DE Human gene signature HUMGS07624.
XX KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
XX KW human; cloning; mapping; non-biased library; diagnosis; detection;
XX KW cell typing; abnormal cell function; ss.
XX OS Homo sapiens.
XX PN W09514772-A1.
XX PD 01-JUN-1995.
XX PF 11-NOV-1994; 94WO-JP01916.
XX PR 12-NOV-1993; 93JP-0355504.
XX PA (MATS/) MATSUBARA K.
XX PA (OKUB/) OKUBO K.
XX PI Matsubara K, Okubo K;
XX DR WPI; 1995-206931/27.
XX PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
XX PT for diagnosis of abnormal cell function, by preparing cDNA that
XX PT reflects relative abundance of corresp. mRNA in specific human
XX PT tissues
XX PS Claim 1; Page 1835; 2245pp; Japanese.
XX CC A single-stranded DNA (or its complementary strand or the corresp.
XX CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
XX CC given in AAT19001-T26837 and which is able to hybridise to part of
XX CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
XX CC sequences were obtained from 3'-directed cDNA libraries prepared
XX CC from various human tissues; synthesis of cDNA was initiated from the
XX CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
XX CC untranslated sequence is unique to a particular mRNA species, almost
XX CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
XX CC is constructed so as to reflect accurately the relative abundance of
XX CC different mRNAs in the particular tissue from which it was derived.
XX CC The appearance frequency of a given GS in a cDNA library can be
XX CC determined (esp. using primers and probes derived from the GS
XX CC sequences) as a means of diagnosing abnormal cell function or for
XX CC recognising different cell types.
XX SQ Sequence 320 BP; 74 A; 75 C; 86 G; 81 T; 4 other;

Query Match 9.6%; Score 313.4; DB 16; Length 320;
Best Local Similarity 98.7%; Pred. No. 4.1e-66;
Matches 314; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2752 GATCAGCAGGGTGCCATTATGTTCTTCTCTAGCCCCCTCAGGAAGAGGACTA 2811
```

```
Db 1 GATCAGCAGGGTGCCATTATGTTCTTCTCTAGCCCCCTCAGGAAGAGGACTA 60
Qy 2812 TATTTGTACTGTACTCCCTAGGGGTTCTGGAAGGGAACATGGAATCAGGATTCATAGAC 2871
Db 61 TATTTGTACTGTACTCCCTAGGGGTTCTGGAAGGGAACATGGAATCAGGATTCATAGAC 120
Qy 2872 TGATAGGCCCTATCCACAAGGGCCATGACTGGGAAAAGGATATGGAGCAGAGGAGATT 2931
Db 121 TGATAGGCCCTATCCACAAGGGCCATGACTGGGAAAAGGATATGGAGCAGAGGAGATT 180
Qy 2932 GGGATTTTAGGGTGACGTACGCTACCCCTAAACTTTTGGTGGCCTGGGGCATCTCTGA 2991
Db 181 GGGATTTNAGGGTGACGTACGCTACCCCTAAACTTTTGGTGGCCTGGGGTATCTCTGA 240
Qy 2992 GGCCAGACTGTTAAGCAGGCTCTGCTGGGCTGTTTACTGCTACCCACCTCTGCACCTGC 3051
Db 241 GGCCAGACTGTTAAGCAGGCTCTGCTGGGCTGTTTACTGCTACCCACCTCTGCACCTGC 300
Qy 3052 TGCTTGAGACTCCATCC 3069
Db 301 TGTCTTGAGACTCCATCC 318

RESULT 13
ABL38341
ID ABL38341 standard; cDNA; 313 BP.
XX AC
XX ABL38341;
XX DT 08-APR-2002 (first entry)
XX DE Human colon tumour antigen polynucleotide SEQ ID NO:1930.
XX KW Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
XX KW colon tumour metastatic antigen; diagnosis; gene; ss.
XX OS Homo sapiens.
XX PN W0200196388-A2.
XX PD 20-DEC-2001.
XX PF 08-JUN-2001; 2001WO-US18557.
XX PR 09-JUN-2000; 2000US-210899P.
XX PR 20-FEB-2001; 2001US-270216P.
XX PA (CORI-) CORIXA CORP.
XX PI Jiang Y, Harlocker SL, Secrlist H;
XX DR WPI; 2002-114514/15.
XX PT Novel isolated colon tumor polynucleotide differentially expressed in
XX PT colon tumor or colon metastatic tumor and polypeptides encoded by them,
XX PT useful for inhibiting development of cancer in patient -
XX PS Claim 1; SEQ ID 1930; 105pp; English.
XX CC ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
XX CC which were isolated from human colon tumour and colon metastatic tumour
XX CC cDNA libraries. (I) have cytostatic activity and can be used in vaccine
XX CC production. (I) can be used for stimulating and/or expanding T cells
XX CC specific for a tumour protein on contact with the T cells. They are also
XX CC useful for inhibiting the development of cancer in a patient. (I) can be
XX CC used as probes or primers for nucleic acid hybridisation, for preparing
XX CC mutant species primers, or primers for use in genetic constructions. (I)
XX CC can be used in the diagnosis of a colon tumour.
XX SQ Sequence 313 BP; 68 A; 85 C; 85 G; 73 T; 2 other;

Query Match 9.4%; Score 307; DB 24; Length 313;
```


Db 379 AACAACTGTGATGATGACAGACAGACAAGTTTCCAAAGNACGGGAGAGAAAGCTGGCC 438
Qy 465 AAGGAGATGCGATGCGACTTCTATGAACAAAGTCCCTGCACCAACCTCAACATTAAGAG 524
Db 439 CTCGACTGGAATCAAGTTTCATGAGACACGCGCGAAGCCCAACATCAATGTGGAAT 498
Qy 525 TCATT 529
Db 499 GCATT 503

RESULT 15

ABQ54410
ID ABQ54410 standard; cDNA; 2411 BP.

XX AC ABQ54410;

XX DT 22-AUG-2002 (first entry)

XX DE Human ovarian antigen HCGMA67 cDNA, SEQ ID NO:290.

XX KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
XX KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
XX KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
XX KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
XX KW inflammatory condition; immune disorder; blood disorder;
XX KW cardiovascular disorder; respiratory disorder; neurological disorder;
XX KW gastrointestinal disorder; urinary system disorder; drug screening;
XX KW gene therapy; chromosome mapping; forensic analysis;
XX KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
XX KW antiinflammatory; gynaecological; reproductive; gene; ss.

XX OS Homo sapiens.

XX PN WO200200677-A1.

XX PD 03-JAN-2002.

XX PF 07-JUN-2001; 2001WO-US18569.

XX PR 07-JUN-2000; 2000US-209467P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Birse CE, Rosen CA;

XX DR WPI; 2002-147878/19.

XX DR P-PSDB; ABP4133.

XX FT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
XX FT useful in the prevention, treatment and diagnosis of cancer (e.g.
XX FT ovarian cancer), immune disorders, cardiovascular disorders and
XX FT neurological diseases -

XX PS Claim 1; SEQ ID NO 290; 2922pp; English.

XX CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
XX CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
XX CC encompasses polypeptides 90% identical and polynucleotides 95% identical
XX CC to the sequences of the invention. The invention additionally relates to
XX CC recombinant vectors and host cells comprising human ovarian antigen
XX CC polynucleotides, antibodies against human ovarian antigens, and the use
XX CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
XX CC treating, prognosing or preventing various ovary and/or breast-related
XX CC disorders. Such conditions include ovarian cancer and breast cancer, and
XX CC metastatic tumours of ovarian or breast origin, reproductive system
XX CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
XX CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
XX CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
XX CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
XX CC vaginitis), immune disorders (e.g., congenital and acquired
XX CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
XX CC blood-related disorders (e.g., anaemia), cardiovascular disorders,

CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents cDNA encoding a human ovarian antigen of the
CC invention.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 2411 BP; 600 A; 671 C; 583 G; 547 T; 10 other;

Query Match 7.0%; Score 228.4; DB 24; Length 2411;
Best Local Similarity 65.9%; Pred. No. 4.3e-45;
Matches 331; Conservative 0; Mismatches 171; Indels 0; Gaps 0;
Qy 28 CCCCGCTGGCCCCAGTCATGGCGAAGCAGTACGATGCTGTCGGGTGCTGCTGATCG 87
Db 27 CCGGGTCGACCCACGCGCTGGCGAAGACCTACGATTACCTGTTCAAGCTGCTGATCG 86
Qy 88 GGGACTCGGGGTGGGCAAGACCTGCTGTCGCGCTTACCAGCAACAGTTCACACT 147
Db 87 GGGACTCGGGGTGGGCAAGACCTGCTGTCGCGCTTACCAGCAACAGTTCACACT 146
Qy 148 CCTCGCACATCTCCACCATCGTCTGACTTTAAGATCAACACCATAGAGCTACACGCA 207
Db 147 CCACCTTATCTCCACCATAGGAATGACTTTAAATTAGGACCATAGAGCTCAGTGCA 206
Qy 208 TCAAAAGTCGGGATACAGATCTGGGACATCTGGCGGCGAGGAGATACACACCATCAAA 267
Db 207 AGAGAATTAAACTGCAGATATGGGACACAGCCGTCAGGAACGGTTTCGGACGATCAAA 266
Qy 268 AGCAGTACTATCGGGGGGCCCGGAGGATATTTTGGTCTATGACATTAGCAGCGCGCT 327
Db 267 CGGCTTACTACAGGGGTGCAATGGGCATCATGCTGGTCTACGACATCAACACGAGAAGT 326
Qy 328 CTTACCAGACATCATGAAGTGGGTGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 387
Db 327 CCTTCGACACATCCGGAACCTGGATTCGCAACATTTAGGAGCGACGCTCTCGACGCTCG 386
Qy 388 AGAAGATCCTTTATTTGGGAATAAGGCTGATGAGGAGCAGAAACGGCAGTGGGAAGAGC 447
Db 387 AAAAGATGATCTCGGGAACAAGTGTGATGTGAATGACAAGAGACAAGTTTCCAAAGAAC 446
Qy 448 AAGGCGAGCAGCTGGCGAAGGAGTATGGCATGGGACTTCTATGAACAACAGTGCCTGCACA 507
Db 447 GGGGAGAAAAGCTGGCCCTCGACTATGGAATCAAGTTTCATGGAGACCGACGCGAAGGCA 506
Qy 508 ACCCTCAACATTTAAAGAGTCATT 529
Db 507 ACATCAATGTGGAAATGCAAT 528

Search completed: June 23, 2003, 13:36:57

Job time : 670 secs

THIS PAGE BLANK (USPTO)

Result No.	Score	Query Match	Length	DB	ID	Description
1	2690.4	82.6	2700	11	BC014511	Homo sapi
2	947.8	29.1	1081	13	BM551913	AGENCOURT
3	856	26.3	957	14	BQ937298	AGENCOURT
4	840.2	25.8	908	14	BQ716774	AGENCOURT
5	810.6	24.9	901	14	BQ220195	AGENCOURT
6	801.6	24.6	924	13	B1118034	B1118034 602867301

Letlicia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-lilisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalius, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 34 Row: j Column: 5
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis

This clone has the following problem: incomplete processing.

FEATURES

source

Location/Qualifiers

1..2700
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4866926"
 /tissue_type="Colon, adenocarcinoma"
 /clone_lib="NIH_MGC_15"
 /lab_host="DH10B-R"
 /note="vector: pOTB7"

BASE COUNT 642 a 749 c 708 g 601 t

ORIGIN

Query Match 82.6%; Score 2690.4; DB 11; Length 2700;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 541 CAGAGCTGGTCTGCAGGCCCATAGGAGGAGCTGGAAGGCTCCGGATGGTCCACGCA 600
 DB 1 CAGAGCTGGTCTGCAGGCCCATAGGAGGAGCTGGAAGGCTCCGGATGGTCCACGCA 60

QY 601 ATGAGTTGGCACTGGCAGAGCTGGAGGAGGAGGAGGCAACCCAGGGCCCGACGAACT 660
 DB 61 ATGAGTTGGCACTGGCAGAGCTGGAGGAGGAGGAGGCAACCCAGGGCCCGACGAACT 120

QY 661 CTTGCAAAACCTGCTGGTGTGAGTCTGTGTGGGGGACCCACAGACACACCCCTCTTCC 720
 DB 121 CTTGCAAAACCTGCTGGTGTGAGTCTGTGTGGGGGACCCACAGACACACCCCTCTTCC 180

QY 721 CTCAGGAGGCCGTGGGCAGACAGGAGCGGGGCTTTGGCTGTGCTGTCTCTCGT 780
 DB 181 CTCAGGAGGCCGTGGGCAGACAGGAGCGGGGCTTTGGCTGTGCTGTCTCTCGT 240

QY 781 GTGATGACCCCTATTGAGTATCAGTAGCCACTACTCCCTGCTGCTGGCCCTGAGAGCGGCT 840
 DB 241 GTGATGACCCCTATTGAGTATCAGTAGCCACTACTCCCTGCTGCTGGCCCTGAGAGCGGCT 300

QY 841 CTGCTGTATCTCAAGCAGCCCTGTCCCGAGCCGCTCCACCCCTGGAGTGTCTCTTCTCA 900
 DB 301 CTGCTGTATCTCAAGCAGCCCTGTCCCGAGCCGCTCCACCCCTGGAGTGTCTCTTCTCA 360

QY 901 GCCTGTTTCCCGCAGCAGCGCTGTCTACGACCCCGCTGAGTGTGCGCAAGCACTGTCTC 960
 DB 361 GCCTGTTTCCCGCAGCAGCGCTGTCTACGACCCCGCTGAGTGTGCGCAAGCACTGTCTC 420

QY 961 ACCATCCCGCAGCAGCAGCAACAGCCAGGGCTGGAGTCCAGGCCACTTTTCAGTGTCTC 1020
 DB 421 ACCATCCCGCAGCAGCAGCAACAGCCAGGGCTGGAGTCCAGGCCACTTTTCAGTGTCTC 480

QY 1021 CTTTCTCGTGTGATCGTGTCTCTTCTGCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTC 1080
 DB 481 CTTTCTCGTGTGATCGTGTCTCTTCTGCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTC 540

QY 1081 TGACCCCTCCCTCCGCTGCTTTCGATCAAGCTCCTCAACCCCGCTCCCGCTGTGT 1140
 DB 541 TGACCCCTCCCTCCGCTGCTTTCGATCAAGCTCCTCAACCCCGCTCCCGCTGTGT 600

QY 1141 CCTGCTGTGTGATCGTGTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 1200
 DB 601 CCTGCTGTGTGATCGTGTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 1740

QY 1201 GCTCGTGGGAGGTTCCACCTTGGATCCAGGAAGAACCTCCACCTGCTGCTGTGGGTG 1260
 DB 661 GCTCGTGGGAGGTTCCACCTTGGATCCAGGAAGAACCTCCACCTGCTGCTGTGGGTG 720

QY 1261 GGCCAAAGGCTACAGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 1320
 DB 721 GGCCAAAGGCTACAGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 780

QY 1321 TGGGCTTGGCTCCCGCAGTGCAGAACTGGAGCATCGAGGTAGGAGGAACACAGCAA 1380
 DB 781 TGGGCTTGGCTCCCGCAGTGCAGAACTGGAGCATCGAGGTAGGAGGAACACAGCAA 840

QY 1381 CCGGGGAGTCTCGAGCCTGGGGCTGCCCTTACCTTCTACCCATTCTCCGACAGAGCTTTG 1440
 DB 841 CCGGGGAGTCTCGAGCCTGGGGCTGCCCTTACCTTCTACCCATTCTCCGACAGAGCTTTG 900

QY 1441 CCCTTGTCTGGTCCCGCCTGCTCTTGGGGAACCTGAGCTCAGAGGCAAGGTCTTCAG 1500
 DB 901 CCCTTGTCTGGTCCCGCCTGCTCTTGGGGAACCTGAGCTCAGAGGCAAGGTCTTCAG 960

QY 1501 AGAAGAAACAAATAGAGGGTGGCAGGGATAGAAAGTCACTCTCTTCTTCTCTCTCTC 1560
 DB 961 AGAAGAAACAAATAGAGGGTGGCAGGGATAGAAAGTCACTCTCTTCTTCTCTCTCTC 1020

QY 1561 TGCAGCATGACACAAATTTCTCTCCACCTGGCTCCCAATTTAAAGATGTGGACCAAGGC 1620
 DB 1021 TGCAGCATGACACAAATTTCTCTCCACCTGGCTCCCAATTTAAAGATGTGGACCAAGGC 1080

QY 1621 CTGTGGGTACTTCCAGGGGCAAGGAGAGCCCTGGGGTCACTGACACTGTCTAGGCCAACCAT 1680
 DB 1081 CTGTGGGTACTTCCAGGGGCAAGGAGAGCCCTGGGGTCACTGACACTGTCTAGGCCAACCAT 1140

QY 1681 GCATCTCACAAGGGGAGCATTTTGGAAATGAAGACTAGCTCTCTATCTATCTAGTTAAGA 1740
 DB 1141 GCATCTCACAAGGGGAGCATTTTGGAAATGAAGACTAGCTCTCTATCTATCTAGTTAAGA 1200

QY 1741 GCAAGGAGAGCTGGCCAGGACAGCAGTTTGCACAGCAGAGGGGAATGTAGCAACAGCA 1800
 DB 1201 GCAAGGAGAGCTGGCCAGGACAGCAGTTTGCACAGCAGAGGGGAATGTAGCAACAGCA 1260

QY 1801 GGGCTCTTAGGCCCTCTTCCATTTCTTAGGTAAGAAGAGCATTTCTCTCAGACTCCCA 1860
 DB 1261 GGGCTCTTAGGCCCTCTTCCATTTCTTAGGTAAGAAGAGCATTTCTCTCAGACTCCCA 1320

QY 1861 GCGGAGGACTGAGCCTTAGCCTTTCAGCAACCAAGGTTCTCTCTGGGACCCCAAGTTATGG 1920
 DB 1321 GCGGAGGACTGAGCCTTAGCCTTTCAGCAACCAAGGTTCTCTCTGGGACCCCAAGTTATGG 1380

QY 1921 GAGAGGGCAAGACTTTCATGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
 DB 1381 GAGAGGGCAAGACTTTCATGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440

QY 1981 CTGTCTCTTTGGCTTTTAAAGCAAAAGCGCTCATCTTGGCTCTTACCTCTCTGATAGGCTT 2040
 DB 1441 CTGTCTCTTTGGCTTTTAAAGCAAAAGCGCTCATCTTGGCTCTTACCTCTCTGATAGGCTT 1500

QY 2041 GAGGTTTGGCAACCAACACACTGTGCTTACAGGTGGAGGAGAGAGACTCTCTCTTCCAGAG 2100
 DB 1501 GAGGTTTGGCAACCAACACACTGTGCTTACAGGTGGAGGAGAGAGACTCTCTCTTCCAGAG 1560

QY 2101 TGCTATGTTCAGGAAGTTTCTTTAAACCCATATGSCCCAGAGTAGCTCTGTAGGAGGCC 2160
 DB 1561 TGCTATGTTCAGGAAGTTTCTTTAAACCCATATGSCCCAGAGTAGCTCTGTAGGAGGCC 1620

QY 2161 TTTAAAGACGGAACAGTAATTTTACCAGTTTCTTACTGGGGTTCCTTGGCCACCGTCCCAAGG 2220
 DB 1621 TTTAAAGACGGAACAGTAATTTTACCAGTTTCTTACTGGGGTTCCTTGGCCACCGTCCCAAGG 1680

QY 2221 TGGGCGAGGCTTAGGAGAGGGTCAATTTTAAAGCCACACATAGCTGACGTGCGTGGGTG 2280
 DB 1681 TGGGCGAGGCTTAGGAGAGGGTCAATTTTAAAGCCACACATAGCTGACGTGCGTGGGTG 1740

```

QY 2281 CAGCCAAACAAAGAACTGGGTGTGAGTATTATCATCACTAAGAACCAAAATCCAGGCA 2340
Db 1741 CAGCCAAACAAAGAACTGGGTGTGAGTATTATCATCACTAAGAACCAAAATCCAGGCA 1800
QY 2341 CTATATGTGAAGGATAAGAACTCACTTCTTACTCTCTCCAAAGAAAGTGGGGAAGA 2400
Db 1801 CTATATGTGAAGGATAAGAACTCACTTCTTACTCTCTCCAAAGAAAGTGGGGAAGA 1860
QY 2401 ACCATCAAACTTTCTCTCTGACTTACCAACACGAGGAAACACAGCAGGAGGCTGCTCA 2460
Db 1861 ACCATCAAACTTTCTCTCTGACTTACCAACACGAGGAAACACAGCAGGAGGCTGCTCA 1920
QY 2461 GGACTTAGGACAGGATATAGCTTAGATGTGGAAGCAAGAGAGCAGCAAGTGTAA 2520
Db 1921 GGACTTAGGACAGGATATAGCTTAGATGTGGAAGCAAGAGAGCAGCAAGTGTAA 1980
QY 2521 ATCACTGGCTTAATGAGAAAGGAGACAGCTTCTAGGATGAAGCTGTCACTAGGCTGG 2580
Db 1981 ATCACTGGCTTAATGAGAAAGGAGACAGCTTCTAGGATGAAGCTGTCACTAGGCTGG 2040
QY 2581 AGTTGCTTCTTGAAGATGGGACTTCTTGGGTATCAAGACCTATGCCACATCACACTGGG 2640
Db 2041 AGTTGCTTCTTGAAGATGGGACTTCTTGGGTATCAAGACCTATGCCACATCACACTGGG 2100
QY 2641 GCTAGGGAAGTAGTGATGCCAGCCTCAAGTCTGTCTTCCAGCCAGGACTTGAAGATT 2700
Db 2101 GCTAGGGAAGTAGTGATGCCAGCCTCAAGTCTGTCTTCCAGCCAGGACTTGAAGATT 2160
QY 2701 ATATTGGGCACTGGCTCAATCTGTGGACCAAGTATTTCAGCTTTCCTGAAGATCAGGCA 2760
Db 2161 ATATTGGGCACTGGCTCAATCTGTGGACCAAGTATTTCAGCTTTCCTGAAGATCAGGCA 2220
QY 2761 GGSTGGCAATTCATTTCTTCTCTCTAGCCCTCAGGAAAGAGCACTATATTGTAC 2820
Db 2221 GGSTGGCAATTCATTTCTTCTCTCTAGCCCTCAGGAAAGAGCACTATATTGTAC 2280
QY 2821 TGTACCTTAGGGTCTTGGAGGAAACATGAAATCAGATTTCTATAGACTATAGGCC 2880
Db 2281 TGTACCTTAGGGTCTTGGAGGAAACATGAAATCAGATTTCTATAGACTATAGGCC 2340
QY 2881 CTATCCACAGGGCCATGACTGGGAAAGTATGGGAGCAGAGAGAAATTTGGATTTA 2940
Db 2341 CTATCCACAGGGCCATGACTGGGAAAGTATGGGAGCAGAGAGAAATTTGGATTTA 2400
QY 2941 GGTGACAGCTACGCTACCCCTAACTTTTGGTGGCTGGGGCATGCTTGAGGCCAGAC 3000
Db 2401 GGTGACAGCTACGCTACCCCTAACTTTTGGTGGCTGGGGCATGCTTGAGGCCAGAC 2460
QY 3001 TGTAAAGCAGGCTCTGCTGGCCCTTTTACTCGTCAACACCTCTGCACTCTCTGTAG 3060
Db 2461 TGTAAAGCAGGCTCTGCTGGCCCTTTTACTCGTCAACACCTCTGCACTCTCTGTAG 2520
QY 3061 ACTCCATCCAGCCCGCAGGACGACCTGCTCTGAGCCCTCCACTATCTCCCTGTGACGG 3120
Db 2521 ACTCCATCCAGCCCGCAGGACGACCTGCTCTGAGCCCTCCACTATCTCCCTGTGACGG 2580
QY 3121 GTGAACCTGTGTACTGTCTCGGCTCCATATATGAATTTGTGAGCAGGGTTTCATCTATT 3180
Db 2581 GTGAACCTGTGTACTGTCTCGGCTCCATATATGAATTTGTGAGCAGGGTTTCATCTATT 2640
QY 3181 TTAACACAGATGTTTACAAAATAAGATTATTTCACACCAACCAAAAAA 3240
Db 2641 TTAACACAGATGTTTACAAAATAAGATTATTTCACACCAACCAAAAAA 2700

```

RESULT 2

```

BM551913 LOCUS 1081 bp mRNA linear EST 20-FEB-2002
DEFINITION AGENCOURT_6575472 NIH_MGC_98 Homo sapiens cdna clone IMAGE:5479433
5', mRNA sequence.
ACCESSION BM551913
VERSION BM551913.1 GI:18789340
KEYWORDS EST.

```

```

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1081)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2000 row: k column: 18
High quality sequence stop: 662.
FEATURES
source
1. 1081
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5479433"
/tissue_type="astrocytoma grade IV, cell line"
/notes="Organ: brain; Vector: pOTB7; Site: 1: XhoI; Site 2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 260 a 295 c 295 g 229 t
ORIGIN
Query Match 29.1%; Score 947.8; DB 13; Length 1081;
Best Local Similarity 95.4%; Pred. No. 1.4e-153;
Matches 1029; Conservative 0; Mismatches 44; Indels 6; Gaps 5;
QY 1178 GCTATCCAAAGGGATGACCCAGGCTCGTGGGAGGTTCCACCTTGGATCCAGAGAA 1237
Db 1 GCTATCCAAAGGGATGACCCAGGCTCGTGGGAGGTTCCACCTTGGATCCAGAGAA 60
QY 1238 CCCTCCACCTGCTCGTGGTGGGCCAAAGGCTACAGGCTGCTTCTTCTTCCCCCA 1297
Db 61 CCCTCCACCTGCTCGTGGTGGGCCAAAGGCTACAGGCTGCTTCTTCTTCCCCCA 120
QY 1298 CCCCCACTGTCCCTCATGTGCCATGGCCCTGCCCTCCCACTGACCTCGGAAAGTGGAGCA 1357
Db 121 CCCCCACTGTCCCTCATGTGCCATGGCCCTGCCCTCCCACTGACCTCGGAAAGTGGAGCA 180
QY 1358 TCAGGTAGGAGGAAACACACACCGGGAGTCTTCAGGCTTGGGGCTGCCTTACCTCTA 1417
Db 181 TCAGGTAGGAGGAAACACACCGGGAGTCTTCAGGCTTGGGGCTGCCTTACCTCTA 240
QY 1418 CCATTTCCCGACACAGCTTTCCTTGTGCTGCTGCCCTGCTGCTTGGGGACT 1477
Db 241 CCATTTCCCGACACAGCTTTCCTTGTGCTGCTGCCCTGCTGCTTGGGGACT 300
QY 1478 GAGCTCAGAGCGAGGCTTTCAGAGAGGAAACAAAATGAGGGTGGCAGGATATAAAG 1537
Db 301 GAGCTCAGAGCGAGGCTTTCAGAGAGGAAACAAAATGAGGGTGGCAGGATATAAAG 360
QY 1538 TCACCTCCATCTCTACCTCCCATGACATGAACAATTTCTTCTCCACTGGCTCCCA 1597
Db 361 TCACCTCCATCTCTACCTCCCATGACATGAACAATTTCTTCTCCACTGGCTCCCA 420
QY 1598 AATTAAAGATGCGACCAAGGCTGTGGGTACTTCCAGGGCAGGAGAGCCCTGGGGTC 1657
Db 421 AATTAAAGATGCGACCAAGGCTGTGGGTACTTCCAGGGCAGGAGAGCCCTGGGGTC 480

```


[illegible]

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC/DCTD/DTF
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM13331 row: e column: 07
 High quality sequence stop: 606.
 Location/Qualifiers
 1..901
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6061422"
 /clone_lib="NIH_MGC_72"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: PCWV-SPO76; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

BASE COUNT 251 a 203 c 242 g 205 t

Query Match 24.9%; Score 810.6; DB 14; Length 901;
 Best Local Similarity 98.7%; Pred. No. 6.2e-130;
 Matches 838; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

2160 CTTTAAAGACGGAACAAAGTAATTTACCACTTCTACTGGGTTCCTGCCACCGTCCCAAG 2219
 |||||||
 1 CTTTAAAGACGGAACAAAGTAATTTACCACTTCTACTGGGTTCCTGCCACCGTCCCAAG 60
 |||||||
 2220 GTGGCGAGGCGCTAGGAAGAGGGTCATCTTAAGCCACACATAGTCGCACATGCGTGGCT 2279
 |||||||
 61 GTGGCGAGGCGCTAGGAAGAGGGTCATCTTAAGCCACACATAGTCGCACATGCGTGGCT 120
 |||||||
 2280 GCAGCCAAACAAAGACACTGGGTGTGAGTATTCATCACTAAGAACCAAAATCCAGGC 2339
 |||||||
 121 GCAGCCAAACAAAGACACTGGGTGTGAGTATTCATCACTAAGAACCAAAATCCAGGC 180
 |||||||
 2340 ACTCATATGTGAAGTAAGAAGACCTCAGCTTCCTTACTCTCCAAAAGAAAGTGGGAAG 2399
 |||||||
 181 ACTCATATGTGAAGTAAGAAGACCTCAGCTTCCTTACTCTCCAAAAGAAAGTGGGAAG 240
 |||||||
 2400 AACCATCAACCTTTCTCTGACITACCAACACGAGAAACAGCAGGAGAGGGTGGCTC 2459
 |||||||
 241 AACCATCAACCTTTCTCTGACITACCAACACGAGAAACAGCAGGAGAGGGTGGCTC 300
 |||||||
 2460 AGGACTTAGGACAGGCTATAGCTTAGATGGTGGAAAGCAAAAGGAGCAGGAAGTTGTA 2519
 |||||||
 301 AGGACTTAGGACAGGCTATAGCTTAGATGGTGGAAAGCAAAAGGAGCAGGAAGTTGTA 360
 |||||||
 2520 AATCACTGGCTAATGAGAAAGAGACAGCTAACTTAGATGAAGCTGTGACTAGGCTG 2579
 |||||||
 361 AATCACTGGCTAATGAGAAAGAGACAGCTAACTTAGATGAAGCTGTGACTAGGCTG 420
 |||||||
 2580 GAGTTGCTTCTTGAGATGGGACCTCTGGGTATCAACACCTATGCCACATCACTACG 2639
 |||||||
 421 GAGTTGCTTCTTGAGATGGGACCTCTGGGTATCAACACCTATGCCACATCACTACG 480
 |||||||
 2640 GGCTAGGGAAGTAGGTGATGCCAGCCCTCAAGTGTGTCTTTCAGCCAGGACCTTGAGAAGT 2699
 |||||||
 481 GGCTAGGGAAGTAGGTGATGCCAGCCCTCAAGTGTGTCTTTCAGCCAGGACCTTGAGAAGT 540
 |||||||
 2700 TATATTGGGAGGTGCTCCAAATCTGTGGACCAAGTATTTCCAGCTTTCCCTTGAAGATCAGGC 2759
 |||||||
 541 TATATTGGGAGGTGCTCCAAATCTGTGGACCAAGTATTTCCAGCTTTCCCTTGAAGATCAGGC 600
 |||||||
 2760 AGGGTCCCATTCATGCTCTTCTCTAGCCCTCCTCAGGAAGGAGGACTATATTGTA 2819
 |||||||
 601 AGGGTCCCATTCATGCTCTTCTCTAGCCCTCCTCAGGAAGGAGGACTATATTGTA 660
 |||||||

QY 2820 CTGTACCCCTAGGGTCTTGGGAAGGAAACATGGAATCAGGATCTCTATAGACTGATAGGC 2879
 |||||||
 DB 661 CTGTACCCCTAGGGTCTTGGGAAGGAAACATGGAATCAGGATCTCTATAGACTGATAGGC 720
 |||||||
 QY 2880 CCTATCCACAAAGGCGCATGACTGGGAAAAGCTATGGGAGCAG-AAGGAGAATTGGGATTT 2938
 |||||||
 DB 721 CCTATCCACAAAGGCGCATGACTGGGAAAAGCTATGGGAGCAGAAAGGAGAATTGGGATTT 780
 |||||||
 QY 2939 TAGGGTCAGCTAGCTCACCTCAAACTTTTGGTGGCCT-GGGGCATCTCTTGAGGCCCA 2997
 |||||||
 DB 781 TAAGGTGAGCTACCTCACCTTAACTTTTGGGGCCTGGGGCATCTCTTGAGGCCCA 840
 |||||||
 QY 2998 GACTCTTAA 3006
 |||||||
 DB 841 GACTGGTTA 849

RESULT 6

BI118034
 LOCUS 602867301F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:5015713 5',
 DEFINITION mRNA sequence.

ACCESSION BI118034
 VERSION BI118034.1 GI:14568935
 KEYWORDS EST.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 924)

NIH-MGC <http://mgs.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LLCM1825 row: b column: 02

High quality sequence stop: 893.

FEATURES

source

1..924
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5015713"
 /clone_lib="NIH_MGC_7"
 /tissue_type="small cell carcinoma"
 /cell_line="MGC3"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 260 a 225 c 244 g 195 t

ORIGIN

Query Match 24.6%; Score 801.6; DB 13; Length 924;
 Best Local Similarity 93.7%; Pred. No. 2.2e-128;
 Matches 868; Conservative 0; Mismatches 54; Indels 4; Gaps 3;
 QY 1504 AGGAACAAATAGAGGGTGGCAGGATAAAAGTCAGCTCCTTCTTACTCCCATGC 1563
 |||||||
 DB 2 AGGAACAAATAGAGGGTGGCAGGATAAAAGTCAGCTCCTTCTTACTCCCATGC 61
 |||||||
 QY 1564 AGCATGAACAAATTTCTCTCCACCTGGCTCCCAAAATTAAGATGTGGACCAAGGCGCTG 1623
 |||||||

Db 62 AGCATGACAAATTTCTCTCCACTGCTCCCAATTTAAAGATGTGGCAAGGCGTG 121
Qy 1624 TGGTACTCCAGGGGCAAGAGAGCCCTGGGGTCACTGACACTGTCCAGCCCAACCATGCA 1683
Db 122 TGGTACTCCAGGGGCAAGAGAGCCCTGGGGTCACTGACACTGTCCAGCCCAACCATGCA 181
Qy 1684 CTCCACAAGGGGAGCAATTTGGAAATGAAGACTAGCTCTATGATATCAGTTTAAGAGCA 1743
Db 182 CTCACAAAGGGGAGCAATTTGGAAATGAAGACTAGCTCTATGATATCAGTTTAAGAGCA 241
Qy 1744 AGGAGAGCTGGCCAGGAGCAGCAGTTTGCACACAGAGGGGAATGTAGCAACAGCAGGG 1803
Db 242 AGGAGAGCTGGCCA-GGACAGAGTTTGCACACAGAGGGGAATGTAGCAACAGCAGGG 300
Qy 1804 CCTCCTAGGCCCCATCTTCCATTTCTTAGTAAAGAGAGCATTTCTCAGACTCCCAAGGC 1863
Db 301 CCTCCTAGGCCCCATCTTCCATTTCTTAGTAAAGAGAGCATTTCTCAGACTCCCAAGGC 360
Qy 1864 GGAGACTAGCCCTAGCCCTCAGCAACAGAGTTCTCTGGGACCCAAAGTTTATGGAG 1923
Db 361 GGAGACTAGCCCTAGCCCTCAGCAACAGAGTTCTCTGGGACCCAAAGTTTATGGAG 420
Qy 1924 AAGGCAAGACTTTCATGGGAAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1983
Db 421 AAGGCAAGACTTTCATGGGAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
Qy 1984 TTCTCTTTGGCTTTAAGACAAAGCGCTCATCTTTGCCCTCTACCTCTGATGAGCTTGAG 2043
Db 481 TTCTCTTTGGCTTTAAGACAAAGCGCTCATCTTTGCCCTCTACCTCTGATGAGCTTGAG 540
Qy 2044 GGTTCGCCAACCACTGTGGCTACAGTGGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 2103
Db 541 GGTTCGCCAACCACTGTGGCTACAGTGGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 600
Qy 2104 TATGTTGAGGAGTTTCTTTAAACCCATATATGCCCCAAGAGTAGCTGCTAGGAGGCCCTT 2163
Db 601 TATGTTGAGGAGTTTCTTTAAACCC--ATATGGCCAGAGTAGTCTGATGAGAGGCCCTT 658
Qy 2164 AAAGCGGAAACAGTAATTTACCAGTTTACTG--GGTTCTCTGCCCAAGGTTCCCAAGGTG 2222
Db 659 AAAGCGGAAACAGTAATTTACCAGTTTACTGAGGGTCCCTGGCCACCGTCCCAAGGTG 718
Qy 2223 GCGAGGCTAGGAGAGGTCATCTTTAAGCCACACATAGCTAGCTAGCTAGCTAGCTAGCT 2282
Db 719 GCGAGGCTAGGAGAGGTCATCTTTAAGCCACACATAGCTAGCTAGCTAGCTAGCTAGCT 778
Qy 2283 GCCAAACAAAGAACTGGGTGTGATGATTTCATCAACTAAGAAACCAAAATCCAGGCACT 2342
Db 779 GCCAAACAGAACTGGGTGTGATGATTTCATCAAGTAAGAAACCAAAATCCAGGCACT 838
Qy 2343 CATATGTGAAGGATGAAGAACTTCTTCTTACTCTTCCCAAAAGAGTGGGAAAGAAC 2402
Db 839 CATATGTGAAGGATGAAGAACTTCTTCTTACTCTTCCCAAAAGAGTGGGAAAGAAC 898
Qy 2403 CATCAAACTTCTCTCTGACTTACC 2428
Db 899 ATTAACCTTACTCTGATTACAAAGC 924

RESULT 7
BQ277739 1046 bp mRNA linear EST 07-MAY-2002
LOCUS AGENCOURT_7047903 NIH_MGC_109 Homo sapiens cdna clone IMAGE:5804597
DEFINITION 5', mRNA sequence.
ACCESSION BQ277739
VERSION BQ277739.1 GI:20487947
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1046)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: ATCC
cdna Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LCM2042 row: h column: 06
High quality sequence stop: 570.

FEATURES
source
Location/Qualifiers
1..1046
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5804597"
/lab_host="NIH_MGC_109"
/tissue_type="teratocarcinoma, cell line"
/note="Organ: ovary; Vector: pOTB7; Site: 1: EcoRI; Site 2:
XhoI; cdna made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT 297 a 250 c 291 g 208 t
ORIGIN

Query Match 24.5%; Score 799.2; DB 14; Length 1046;
Best Local Similarity 96.9%; Pred. No. 5.4e-128;
Matches 847; Conservative 0; Mismatches 23; Indels 4; Gaps 3;
Qy 1547 TTCCTACCTCCCATGAGCATGACACAAATTTCTCTCCACTGGCTCCCAAAATTTAAAG 1606
Db 1 TTCCTACCTCCCATGAGCATGACACAAATTTCTCTCCACTGGCTCCCAAAATTTAAAG 60
Qy 1607 ATGTGGACCAAGGCTGTGGTACTCCAGGGGAGAGAGCCCTGGGTCAGTCACACT 1666
Db 61 ATGTGGACCAAGGCTGTGGTACTCCAGGGGAGAGAGCCCTGGGTCAGTCACACT 120
Qy 1667 GTCAGGCCAACCATGCACTCCACAAAGGGGAGCATTTTGGAAATGAAGACTAGCTCTAT 1726
Db 121 GTCAGGCCAACCATGCACTCCACAAAGGGGAGCATTTTGGAAATGAAGACTAGCTCTAT 180
Qy 1727 GTATCAGTTTAAGAGCAAGGAGAGCTGGCCAGGGACAGCATGTTTGCACAGCAGAGGGA 1786
Db 181 GTATCAGTTTAAGAGCAAGGAGAGCTGGCCAGGGACAGCATGTTTGCACAGCAGAGGGA 240
Qy 1787 ATGTAGCAACAGCAGGCGCTCTAGGCCCATCTTCCATTTCTTAGTAAAGAGCAATT 1846
Db 241 ATGTAGCAACAGCAGGCGCTCTAGGCCCATCTTCCATTTCTTAGTAAAGAGCAATT 300
Qy 1847 TCCTCAGACTCCAGGCGGAGGACTGAGCCTAGCCTTCCAGCAACCAAGGTTCTCTCTGGA 1906
Db 301 TCCTCAGACTCCAGGCGGAGGACTGAGCCTAGCCTTCCAGCAACCAAGGTTCTCTCTGGA 360
Qy 1907 CCCAAAGTTTATGGGAGAGGCAAGACTTTCATGGGAGAGAGAGAGGAGGAGGAGGAGGAG 1966
Db 361 CCCAAAGTTTATGGGAGAGGCAAGACTTTCATGGGAGAGAGAGAGGAGGAGGAGGAGGAG 420
Qy 1967 AGAAACGCTTGGTGTCTCTTTGGCTTTAAGACAAAGGCTCATCTTGGCCCTCTAC 2026
Db 421 ACAACGCTTGGTGTCTCTTTGGCTTTAAGACAAAGGCTCATCTTGGCCCTCTAC 480
Qy 2027 CTCCTGATAGCTTGGGTTTGGCAACCACTGCTAGCTAGGAGGAGGAGGAGGAGGAGGAG 2086
Db 481 CTCCTGATAGCTTGGGTTTGGCAACCACTGCTAGCTAGGAGGAGGAGGAGGAGGAGGAG 540


```

QY 2087 TCCTCTCCAGAGTGTATGTTTCAGGAGTTTCTTTAAACCCCATATATGCCCAAGAGTAG 2146
Db 541 TCCTTCTCCAGAGTGTATGTTTCAGGAGTTTCTTTAAACCCCATATATGCCCAAGAGTAG 600
QY 2147 CTCGTAGGAGGCCCTTTAAAGACGGAACAGTAATTTTACCAGTTCTACTGGGGTTCCTGCG 2206
Db 601 CTCCTATGAGGCCCTTTAAAGACGGAACAGTAATTTTACCAGTTCTACTGGGGTTCCTGCG 660
QY 2207 CCACCGTCCCAAGGTGGGGAGGCCCTAGGAAGAGGTGCTTCTTAAGCCACACATATAGCT 2266
Db 661 CCACCGTCCCAAGGTGGGGAGGCCCTATGAAGAGGTGCTTCTTAAGCCACACATATAGCT 720
QY 2267 GCACGTGGGTGGTGCAGCCCAACAAAGAACTGGGTGTGAGTATTCATCACTAAGAAC 2326
Db 721 GCACGTGGGTGGTGCAGCCCAACAAAGAACTGGGTGTGAGTATTCATCACTAAGAAC 780
QY 2327 CAAATCCAGGCGCACTCATATCT-GAAGGATAGAAC--TCACCTTCTTACTCTCCAA 2383
Db 781 CCAATCCAGGCGCACTCATATCTGGAGGATAGAACCGTCACTTCTTACTCTCGCAA 840
QY 2384 AAAGAAGTGGGG-AAAGAACCATCAACCTTTCC 2416
Db 841 AAAAAGGGGGAAAGAACCATCGAACTTTTC 874

RESULT 8
BM476862
LOCUS
DEFINITION
AGENCOURT_5481739 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5555414
5', mRNA sequence.
ACCESSION
BM476862
VERSION
EST.
KEYWORDS
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1027)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: AFCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M12275 row: i column: 15
High quality sequence stop: 607.
Location/Qualifiers
1. .1027
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5555414"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."
BASE COUNT 260 a 272 c 283 g 211 t 1 others
ORIGIN
Query Match 24.0%; Score 781.2; DB 13; Length 1027;
Best Local Similarity 97.9%; Pred. No. 6.8e-125;
Matches. 823; Conservative 0; Mismatches 14; Indels 4; Gaps 3;
QY 1354 AGCATCAGGTAGGAGGAAACAGCAACGGGAGTCTCTGAGCCCTGGGCTGCCCTACC 1413
Db 2 AGCATCAGGTAGGAGGAAACAGCAACGGGAGTCTCTGAGCCCTGGGCTGCCCTACC 61

```

```

QY 1414 TCTACCCATTCGCCGACAGAGCTTTGCCCTTGGCTTGGCTGCCGCTCTCTTTGGGG 1473
Db 62 TCTACCCATTCGCCGACAGAGCTTTGCCCTTGGCTTGGCTGCCGCTCTCTTTGGGG 121
QY 1474 AACTGAGCTCAGAGCAGGTGCTTCAGAGAGGAACAAAATAGAGGGTGGCAGGATAA 1533
Db 122 AACTGAGCTCAGAGCAGGTGCTTCAGAGAGGAACAAAATAGAGGGTGGCAGGATAA 181
QY 1534 AAAGTCACCTCATTCTCTACTCCCATGCAGATGAACAAATTTCTCTCACTGGCT 1593
Db 182 AAAGTCACCTCATTCTCTACTCCCATGCAGATGAACAAATTTCTCTCACTGGCT 241
QY 1594 CCAAAATTTAAAGATGTGGACCAAGCCCTGTGGGTACTCCAGGGGCAAGAGAGCCCTGG 1653
Db 242 CCAAAATTTAAAGATGTGGACCAAGCCCTGTGGGTACTCCAGGGGCAAGAGAGCCCTGG 301
QY 1654 GGTGAGTGACACTGTGCAGCCCAACCATGCATCCCAAGAGGGGAGCATTTGGAATGAAG 1713
Db 302 GGTGAGTGACACTGTGCAGCCCAACCATGCATCCCAAGAGGGGAGCATTTGGAATGAAG 361
QY 1714 GACTAGCTCCTATGTATCAGGTTAAGAGCAAGGAGAGCTGGCCAGGAGCAGTGTGC 1773
Db 362 GACTAGCTCCTATGTATCAGGTTAAGAGCAAGGAGAGCTGGCCAGGAGCAGTGTGC 421
QY 1774 ACAGCAGAGGGGAATGTAGCAACAGCAGGGGCTCTTAGGCCCATCTTCCATTTCTTAGG 1833
Db 422 ACAGCAGAGGGGAATGTAGCAACAGCAGGGGCTCTTAGGCCCATCTTCCATTTCTTAGG 481
QY 1834 TAAGAAGAGCATTTCTCAGACTCCCGAGGAGCAGTGCCTAGCCTTCAGCAACCAA 1893
Db 482 TAAGAAGAGCATTTCTCAGACTCCCGAGGAGGAGCTGAGCCTTAGCCTTCAGCAACCAA 541
QY 1894 GGTTCCTCTGGGACCCAAAGTTTATGGGAGAGGGCAAGAGCTTCATGGGAGAGAGAG 1953
Db 542 GGTTCCTCTGGGACCCAAAGTTTATGGGAGAGGGCAAGAGCTTCATGGGAGAGAGAG 601
QY 1954 GAAGGCCCTGGGTAGAACGCTTGTGCTGTCTTCTTTGGGCTTTAAGACAAAGCGCTCA 2013
Db 602 GAAGGCCCTGGGTAGANAGCTTGTGCTGTCTTCTTTGGGCTTTAAGACAAAGCGCTCA 661
QY 2014 TCTTGGCCCTCTACCTCTCTGATAGGCTTGAAGGCTTTGCCAACACACACTGTGGCTACAGGT 2072
Db 662 TCTTGGCCCTCTACCTCTCTGATAGGCTTGAAGGCTTTGCCAACACACACTGTGGCTACAGGT 721
QY 2073 GGAGGAGAGAGACTCTCTTCTCCAGAGTGTCTATGTTCAGGAGTCTTTTAAACCCATA 2132
Db 722 GGAGGAGAGAGACTCTCTTCTCCAGAGTGTCTATGTTCAGGAGTCTTTTAAACCCATA 781
QY 2133 TGG-CCCAAGAGTACTCTCTAGGAGGCCCTTT--AAAGACGCAACAGTAATTTACCAGT 2189
Db 782 TGGCCCCAGGAGTAGCTCGTAGAGGCCCTTTTAAAGACGCAACAGTAATTTACCAGT 841
QY 2190 T 2190
Db 842 T 842

RESULT 9
BG676934
LOCUS
DEFINITION
602623549F1 NCI_CCAP_Skn4 Homo sapiens cDNA clone IMAGE:4748437 5',
mRNA sequence.
ACCESSION
BG676934
VERSION
BG676934.1 GI:13908331
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 878)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)

```


.....


```
/clone="IMAGE:4873185"
/clone_lib="NIH_MGC_49"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/notes="organ: skin; Vector: pOTB7; Site:1: XhoI; Site:2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-CDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT      190 a 192 c 220 g 147 t
ORIGIN
Query Match      22.2%; Score 721.8; DB 12; Length 749;
Best Local Similarity 99.5%; Pred. No. 1.3e-114;
Matches 745; Conservative 0; Mismatches 2; Indels 2; Gaps 2;
QY 1339 GACCTGCCAAGTGGAGCATCGAGTAGAGGGGAACAGCAACCGGGAGTCTCGAGCC 1398
Db 1 GACCTGCCAAGTGGAGCATCGAGTAGAGGGGAACAGCAACCGGGAGTCTCGAGCC 60
QY 1399 TGGGGCTGCCCTACCTCTACCATCCCGGACGAGCTTTGCCCTTGGCTGGCCG 1458
Db 61 TGGGGCTGCCCTACCTCTACCATCCCGGACGAGCTTTGCCCTTGGCTGGCCG 120
QY 1459 CTTGCTCTTTGGGGAAGTGAAGTTCAGAGGAGGAGTCTTCAGAGGAAGAAACAAATGAG 1518
Db 121 CTTGCTCTTTGGGGAAGTGAAGTTCAGAGGAGGAGTCTTCAGAGGAAGAAACAAATGAG 180
QY 1519 GGGTGGCAGGGATAAAGTACCTTCTTACCTCCCATGAGCATGAACACAATT 1578
Db 181 GGGTGGCAGGGATAAAGTACCTTCTTACCTCCCATGAGCATGAACACAATT 240
QY 1579 TCTCTCCACCTGGCTCCCAATTTAAAGATGGACCAAGCGCTTGGTACTCCAGGG 1638
Db 241 TCTCTCCACCTGGCTCCCAATTTAAAGATGGACCAAGCGCTTGGTACTCCAGGG 300
QY 1639 CAAGGAGAGCCCTGGGGTCACTGACATGTCTAGGCGCAACCACTGCAACAAAGGGAG 1698
Db 301 CAAGGAGAGCCCTGGGGTCACTGACATGTCTAGGCGCAACCACTGCAACAAAGGGAG 360
QY 1699 CATTGGGAATGAAGACTAGTCTCTATGTATATAGTTAAGCAAGGAGAGCTGGCCA 1758
Db 361 CATTGGGAATGAAGACTAGTCTCTATGTATATAGTTAAGCAAGGAGAGCTGGCCA 420
QY 1759 GGGCAGCAGTTTGCACAGCAGAGGGGAATGTAGCAACAGCAGGCGCTCTAGGCCCAT 1818
Db 421 GGGCAGCAGTTTGCACAGCAGAGGGGAATGTAGCAACAGCAGGCGCTCTAGGCCCAT 480
QY 1819 CTTCCATTTCTTAGCTAAGAGAGCATTTCTCTCAGACTCCCAAGGGAGAGCTAGACCTA 1878
Db 481 CTTCCATTTCTTAGCTAAGAGAGCATTTCTCTCAGACTCCCAAGGGAGAGCTAGACCTA 540
QY 1879 GCCTTCAGCAACCAAGTTCTCTCTGGGACCCCAAGTTTATGGGAAGAGGCAAGACTTC 1938
Db 541 GCCTTCAGCAACCAAGTTCTCTCTGGGACCCCAAGTTTATGGGAAGAGGCAAGACTTC 600
QY 1939 ATGGGAAGAGAGAA -GGAAGGCCCTGGGTAGAAAGCTTGGT -GCTGTCTCTTTGGCT 1996
Db 601 ATGGGAAGAGAGAAAGGAGGCCCTGGGTAGAAAGCTTGGTGGTGGTCTCTTTGGCT 660
QY 1997 TTAAGCAAAAGGCTCATCTTCCCTCTACCTCTGATAGGCTTGGAGGTTTGGCAACCA 2056
Db 661 TTAAGCAAAAGGCTCATCTTCCCTCTACCTCTGATAGGCTTGGAGGTTTGGCAACCA 720
QY 2057 CACTGTGGCTACAGTGGAGGGAAGAGGA 2085
Db 721 CACTGTGGCTACAGTGGAGGGAAGAGGA 749
```

```
RESULT 15
BM049329
LOCUS
DEFINITION
603626279F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:5452987 5',
mRNA sequence.
766 bp mRNA linear EST 07-NOV-2001
BM049329
ACCESSION
BM049329
VERSION
BM049329.1 GI:16778596
KEYWORDS
EST...
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 766)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTB/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1945 row: m column: 20
High quality sequence stop: 758.
Location/Qualifiers
1..766
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5452987"
/tissue_lib="NIH_MGC_40"
/lab_host="DH10B (phage-resistant)"
/notes="organ: prostate; Vector: pOTB7; Site:1: XhoI;
Site:2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGCAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT      190 a 202 c 221 g 153 t
ORIGIN
Query Match      22.0%; Score 716.8; DB 13; Length 766;
Best Local Similarity 99.2%; Pred. No. 9.3e-114;
Matches 762; Conservative 0; Mismatches 2; Indels 4; Gaps 4;
QY 1312 CATGTGCCATGGCGCTGCCCTCCCGAGTGCACCTGCGAAGTGGAGCATCGAGTAGGAGG 1371
Db 2 CATGTGCCATGGCGCTGCCCTCCCGAGTGCACCTGCGAAGTGGAGCATCGAGTAGGAGG 61
QY 1372 AAACAGCAACCGGGGAGTCTCGAGCGCTGGCGTGCCTTACCTTACCATTCCTCCCGACC 1431
Db 62 AAACAGCAACCGGGGAGTCTCGAGCGCTGGCGTGCCTTACCTTACCATTCCTCCCGACC 121
QY 1432 AGAGCTTTGCCCTTGGCTGGCGCTGCCCTTCTTTGGGAACCTAGCTCAGAGGCAG 1491
Db 122 AGAGCTTTGCCCTTGGCTGGCGCTGCCCTTCTTTGGGAACCTAGCTCAGAGGCAG 181
QY 1492 GTGCTTCAGAGAAGGAACAAATGAGGGGTGGCAGGATATAAAAGTACCTTCATCTCTC 1551
Db 182 GTGCTTCAGAGAAGGAACAAATGAGGGGTGGCAGGATATAAAAGTACCTTCATCTCTC 241
QY 1552 TACCTCCCATGAGCATGAACACAATTCTCTCCAGCTGGCTCCCAAAATTAAGATGTG 1611
Db 242 TACCTCCCATGAGCATGAACACAATTCTCTCCAGCTGGCTCCCAAAATTAAGATGTG 301
QY 1612 GACCAAGGCTCTGGGTACTCCAGGGGCAAGGAGCCCTGGGTGAGTCACTGACTCTCAG 1671
Db 302 GACCAAGGCTCTGGGTACTCCCA -GGGCAAGGAGAGCCCTGGGTGAGTCACTGACTCTCAG 360
```

1672	QY	GCACACCATGCACCTCCACAAAGGGGAGACATTTGGAAATGAAGGACTAGCTCCTATGTATC	1731
361	Db	GCACACCATGCACCTCCACAAAGGGGAGACATTTGGAAATGAAGGACTAGCTCCTATGTATC	420
1732	QY	AGGTTTAAGACGAAGGAGAGCTGGCCAGGACACAGCATTTTGACACAGCAGAGGGGAATGTA	1791
421	Db	AGGTTTAAGACGAAGGAGAGCTGGCCAGGACACAGCATTTTGACACAGCAGAGGGGAATGTA	480
1792	QY	GCAACAGCAGGGCCCTCTAGGCCCCATCTTCCATTTCTTAGGTAAGAAGAGCATTTCCCTC	1851
481	Db	GCAACAGCAGGGCCCTCTAGGCCCCATCTTCCATTTCTTAGGTAAGAAGAGCATTTCCCTC	540
1852	QY	AGACTCCCAGGCGGAGGACTAGCCTTAGCAACCAAGTTCTCCTGGGACCCAA	1911
541	Db	AGACTCCCAGGCGGAGGACTAGCCTTAGCAACCAAGTTCTCCTGGGACCCAA	600
1912	QY	AGTTTATGGGAAGGGCAAGACATTCATGGGAAGAGCAAGGAGCCCTGGGTAGAAA	1971
601	Db	AGTTTATGGGAAGGGCAAGACATTCATGGGAAGAGCAAGGAGG-CCTGGGTAGAAA	659
1972	QY	CGCTTGGTCTGTCTCTTTGGCCTTTAAGACAAAGCGCTCATCTTGGCCTCTACCTCCT	2031
660	Db	CGCTTGGTCTGTCTCTTTGGCCTTTAAGACAAAGCGCTCATCTTGGCCTCTA-CTCCT	718
2032	QY	GATAGCCTTGAGGGTTTGGCAACCAACACATGTGCTACAGGTGGAGGA	2079
719	Db	GATAGCCTTGAGGG-TTGGCAACCAACATGTGCTACAGGTGGAGGA	765

Search completed: June 23, 2003, 17:08:10
Job time : 4320 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	156.6	4.8	1340	2	US-08-824-873-2	Sequence 2, Appli
2	156.6	4.8	1340	3	US-09-198-181-2	Sequence 2, Appli
3	156	4.8	925	3	US-08-916-901-4	Sequence 4, Appli
4	156	4.8	925	4	US-09-154-602-4	Sequence 4, Appli
5	136.2	4.2	875	4	US-09-075-454-10	Sequence 10, Appli
6	136.2	4.2	2612	4	US-09-484-970B-142	Sequence 142, Appli
7	120.6	3.7	847	2	US-08-773-423-4	Sequence 4, Appli
8	113	3.5	639	4	US-09-399-913-66	Sequence 66, Appli
9	110	3.4	970	3	US-08-888-077A-28	Sequence 28, Appli
10	91.2	2.8	820	3	US-08-741-411-6	Sequence 6, Appli
11	85.8	2.6	7218	1	US-08-232-463-14	Sequence 14, Appli
12	80.6	2.5	890	3	US-08-741-411-4	Sequence 4, Appli
13	80	2.5	1172	4	US-09-075-454-8	Sequence 8, Appli
14	80	2.5	1255	2	US-08-766-551-6	Sequence 6, Appli
15	80	2.5	1533	4	US-09-075-454-11	Sequence 11, Appli
16	79	2.4	848	3	US-08-741-411-2	Sequence 2, Appli
17	77	2.4	1749	4	US-09-149-476-54	Sequence 54, Appli
18	75.6	2.3	1407	4	US-09-493-914-1	Sequence 1, Appli
19	73.6	2.3	803	4	US-09-075-454-13	Sequence 13, Appli
20	71.6	2.2	591	2	US-09-156-979-1	Sequence 1, Appli
21	71.6	2.2	591	4	US-09-387-341-68	Sequence 68, Appli
22	67.2	2.1	702	3	US-08-842-976-2	Sequence 2, Appli
23	67.2	2.1	702	3	US-09-213-397-2	Sequence 2, Appli
24	67.2	2.1	702	3	US-09-416-489-2	Sequence 2, Appli
25	66.2	2.0	331	4	US-09-325-932A-33	Sequence 33, Appli
26	64.8	2.0	615	1	US-08-247-946A-5	Sequence 5, Appli
27	64.8	2.0	615	5	PCN-US95-06420-5	Sequence 5, Appli


```
QY 57 TACGATGCTCTTCCGGCTGCTGATCGGGGACTCCGGGTGGGCAAGACCTGCTG 116
Db 22 TACGAGCTGCGCTTCAAGGTCTGCTGTTGGGGTGGGCAAGACCTGCTG 81
QY 117 CTGTGC-GCCTTCACCGCAACAGCTTCCACTCCCGCACATCTCCACATCGGTG--TT 173
Db 82 CTGGGTGGGATTCAGGATGGTCTTCTCGCGGGGACCTTCATCTCCACGTAAGCAT 141
QY 174 GACTTTAAGATGAAGACCATAGAGTAGACGGCATCAAAGTGGCGGATACAGATCTGGGAC 233
Db 142 GACTTCCGGAAACAAAGTTCTTGACGTGGATGGTGAAGTGAAGCTGACATGTGGGAC 201
QY 234 ACTGACGGCAGGAGAGATACAGACCATCAAGAGAGTACTATCGGGGGCCCGAGGG 293
Db 202 ACAGCTGTGACAGCGGTTCGCGAGTGTACCCATGCTTACCGGATGCTCATGCT 261
QY 294 ATATTTTGGTCTATGACATAGCAGCGAGCGCTCTTACCAGCATCATGAAGTGGGTC 353
Db 262 CTGCTGCTGCTACGATGTCCACCAAGAGCGCTCTTTGACAAACATCCAGGCTGGGTC 321
QY 354 AGTGACGTGGATGACTACGACCAAGCGCTCCAGAGATCCTTATTTGGGATTAAGCT 413
Db 322 ACCGAGATCCAGGATACGCCAGCAGCGTGGCGCTCATGCTGCTGGGGAACAAAGTG 381
QY 414 GATGAGGAGCAAAACGGCAGTGGGAAGAGAGCAAGGCGCAGCTGGCGAAGGATAT 473
Db 382 GACTTCCGCAATGAGCGTGTGTGAAGAGGAGGAGCGGAGAACTGGCCAAAGATAT 441
QY 474 GCGATGGACTTTATGAACAAAGTGCCTGCGACCAACCTCAACATTAAGAGTCAATTCAG 533
Db 442 GGACTGCCCTTATGAGACCAAGCGGAGGAGCGGCGCTCAAGTGGACTTGGCCCTTACA 501
QY 534 CGTCTGACAGACTGCTGCTGCGAGCGCCCATAGGAAGG 570
Db 502 GCCATAGCAAGAGGTGAAGCAGCGCTCCATGAAGG 538
```

RESULT 2

```
US-09-198-184-2
; Sequence 2, Application US/09198184
; Patent No. 6010859
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Guebler, Karl
; TITLE OF INVENTION: NOVEL RAB PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09198,184
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/824,873
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0240 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 2:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1340 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PANCNOT04
; CLONE: 738957
; US-09-198-184-2

Query Match 4.8%; Score 156.6; DB 3; Length 1340;
Best Local Similarity 59.0%; Pred No. 8.1e-30;
Matches 305; Conservative 0; Mismatches 209; Indels 3; Gaps 2;

QY 57 TACGATGCTGTTCCGGCTGCTGATCGGGGACTCCGGGTGGGCAAGACCTGCTG 116
Db 22 TACGAGCTGCGCTTCAAGGTCTGCTGTTGGGGTGGGCAAGACCTGCTG 81
QY 117 CTGTGC-GCCTTCACCGCAACAGGTTCCACTCCCTCGCACATCTCCACATCGGTG--TT 173
Db 82 CTGGGTGGGATTCAGGATGGTCTTCTCGCGGGGACCTTCATCTCCACGTAAGCAT 141
QY 174 GACTTTAAGATGAAGACCATAGAGTAGACGGCATCAAAGTGGCGGATACAGATCTGGGAC 233
Db 142 GACTTCCGGAAACAAAGTTCTTGACGTGGATGGTGAAGTGAAGCTGACATGTGGGAC 201
QY 234 ACTGACGGCAGGAGAGATACAGACCATCAAGAGAGTACTATCGGGGGCCCGAGGG 293
Db 202 ACAGCTGTGACAGCGGTTCGCGAGTGTACCCATGCTTACCGGATGCTCATGCT 261
QY 294 ATATTTTGGTCTATGACATAGCAGCGAGCGCTCTTACCAGCATCATGAAGTGGGTC 353
Db 262 CTGCTGCTGCTACGATGTCCACCAAGAGCGCTCTTTGACAAACATCCAGGCTGGGTC 321
QY 354 AGTGACGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 413
Db 322 ACCGAGATCCAGGATACGCCAGCAGCGTGGCGCTCATGCTGCTGGGGAACAAAGTG 381
QY 414 GATGAGGAGCAAAACGGCAGTGGGAAGAGAGCAAGGCGCAGCTGGCGAAGGATAT 473
Db 382 GACTTCCGCAATGAGCGTGTGTGAAGAGGAGGAGCGGAGAACTGGCCAAAGATAT 441
QY 474 GCGATGGACTTTCTATGAACAAAGTGCCTGCGACCAACCTCAACATTAAGAGTCAATTCAG 533
Db 442 GGACTGCCCTTATGAGACCAAGCGGAGGAGCGGCGCTCAAGTGGACTTGGCCCTTACA 501
QY 534 CGTCTGACAGACTGCTGCTGCGAGCGCCCATAGGAAGG 570
Db 502 GCCATAGCAAGAGGTGAAGCAGCGCTCCATGAAGG 538
```

RESULT 3

```
US-08-916-901-4
; Sequence 4, Application US/08916901
; Patent No. 5892012
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: RAB PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
```

```

; SOFTWARE: FastSeq for Windows Version 2.0
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: US/08/916,901
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
;
; APPLICATION NUMBER:
; FILING DATE:
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0367 US
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
;
; INFORMATION FOR SEQ ID NO: 4:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 925 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; IMMEDIATE SOURCE:
; LIBRARY: LIVRUT04
; CLONE: 2514506
;
; US-08-916-901-4

```

Query Match	4.98;	Score 156;	DB 2;	Length 925;
Best Local Similarity	56.7%;	Pred. No. 9.7e-30;		
Matches	288;	Conservative 0;	Mismatches 220;	Indels 0; Gaps 0;
QY	36	GC CCCAGTCATGCGGAGCAGTAGCAGTGTCTCTTCGGCTGCTGCTGATCGGGGACTCC	95	
Db	57	GC CGCCGCGCATGAACCCGGAATATGACTCTCTTTAAAGCTGCTTTTGAATGGCGACTCA	116	
QY	96	GGGGTGGGGAAGACCTCCCTGCTGTGTGCGGCTTCACCGACAACAGATTTCCACTCTCTCGCAC	155	
Db	117	GGCGTGGGCAAGTCATGCGCTGCTCTCGGCTTTCCTGATCACAGTACACAGAGAGCTAC	176	
QY	156	ATCTCCACCATCCGGTGTGTGACTTTAGATGAAGACCATTAGAGGTAGACGGCATCAAAAGTG	215	
Db	177	ATCAGACACCATCCGGGTGGACTTCAAGATCCGAACCATCGAGCTGGATGGCAAAACTATC	236	
QY	216	CGGATACAGATCTGGACACTGCAGGCGGAGGAGATACCAGACCATCATCAAAAGCAGTAC	275	
Db	237	AAACTTCAGATCTGGGACACAGCGGGCCAGGAACGGTTCGGACCATCACTTCCAGCTAC	296	
QY	276	TATCGCGGGGCCAGGGGATATTTTGGTCTATGACATTAGCGAGCGCTCTTTACCAG	335	
Db	297	TACCGGGGGCTCATGSCATCATCGTGGTCTATGCGTCACTGACCAGGAATCCTAGGCC	356	
QY	336	CACATCATGAGTGGGTGAGTGAGTGGATGAGTCAGCACACGAAGCGGTCCAGAGATC	395	
Db	357	AACGTGAAGCAGTGGGTTCAGAGAGATTTGACCGCTATGCCAGCGAAGACGTCATTAAGCTC	416	
QY	396	CTTATTGGGAATAAGCTGATAGGAGCAGCAACCGCAGCTGGGAAGAGCAGGAGGCCAG	455	
Db	417	CTGGTGGCCACAGAGCGACCTCACCAACAGAGGTGGTGACACACACCCACAGCCAG	476	
QY	456	CAGCTCGCGAAGGAGTATGGCATGGACTTCTATGAACAAGTGCCTGCACCAACCTCAAC	515	
Db	477	GAGTTTGCAGACTCTCTGGGCATCCCTCTTGGAGACGAGCGCCAGAGATGCCACCAAT	536	
QY	516	ATTAAGAGTCATTACGGCGTCTGACAG	543	
Db	537	GTCCGACGACGGCTTCATGACCATGGGTG	564	

RESULT 4
US-09-154-602-4
; Sequence 4, Application US/09154602
; Patent No. 6300472
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti

APPLICANT: Corley, Neil C.
 APPLICANT: Shah, Purvi
 TITLE OF INVENTION: RAB PROTEINS
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Dr.
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/154,602
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/916,901
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0367 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 925 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: LIVRUT04
 CLONE: 2514506
 US-09-154-602-4

[illegible]

Db 477 GAGTTTGAGACTCTCTGGGCATCCCTCTTGGAGACGAGCGCAAGAATGCCACCAAT 536
QY 516 ATTAAGAGTCAATCAGCGCTCGACAG 543
Db 537 GTCGAGCAGCGGCTCATGACCATGGCTG 564

RESULT 5

US-09-075-454-10
; Sequence 10, Application US/09075454
; Patent No. 6391580
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Tang, Y. Tom
; APPLICANT: Lal, Preeti
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Patterson, Chandra
; APPLICANT: Batra, Sajeev
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: RAS PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 6.1/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/075,454
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/766,551
; FILING DATE: DECEMBER 12, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Certone, Michael C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0168-1 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 875 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: UCMCLST01
; CLONE: 1528559
; US-09-075-454-10

Query Match 4.2%; Score 136.2; DB 4; Length 875;
Best Local Similarity 57.5%; Pred. No. 8.9e-25;
Matches 265; Conservative 0; Mismatches 193; Indels 3; Gaps 1;
QY 57 TAGCATGTGCTGTCCGGCTGCTGTGATCGGGGACTCCGGGTGGGCAAGACCTGCGCTG 116
Db 72 TACGACCTCAGCGGCAAGGTGATGCTTCTGGGAGACACAGCGGTGCGCAAAACATGTTTC 131
QY 117 CTGTGCGGCTTACCCACACACGAGTTCGACTCCT---CGCACATCTCCACCATCGGTGT 173
Db 132 CTGATCCCAATTCAAGACGCGGCTTCTGTCCGGAAACCTTCATAGCCACCGTTCGGCATA 191

QY 174 GACTTTAAAGATGAAGACCATTAGAGGTAGACGGGATCAAACTCGGGATACAGATCTGGGAC 233
Db 192 GACTTTCAGGAACCAAGGTGGTGAAGTGGGATGGAGTGAAGCTGCAGATCTGGGAC 251
QY 234 ACTGAGGCGGAGAGATACACAGCATCACAAAGCAGTACTATCGGCGGGCCAGGGG 293
Db 252 ACCGCTGGGAGGAAACGGTCCGAGCGTCACCCATGCTTATTACAGAGATGTCAGGCC 311
QY 294 ATATTTTGGTCTATGACATTAGCAGCGCGCTCTTACCAGCACATCATGAAGTGGGTC 353
Db 312 TTGCTTCTGCTGATGACATCACCAAAATCTCTTTGACACATCATCAGGCGCTGGCTC 371
QY 354 AGTCACGTGATGAGTACGACCAAGAGGCTCCAGAGATCCTTATTGGGAATAGGCT 413
Db 372 ACTGAGATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 431
QY 414 GATGAGGAGCAGAAACGCGAGTGGGAGAGAGAGGCGAGCAGCAGCAGCAGCAGCAGCAG 473
Db 432 GATATGAGCAGCGAAGAGTATCCGTTCCGAGACGAGAGACCTTGGCCAGGAGTAC 491
QY 474 GGCATGGACTTCTATGAACAAGTGCCTGCACCAACCTCAA 514
Db 492 GGTGTTCCCTTCTCTGGAGACGAGCGCAAGACTGGCATGAA 532

RESULT 6

US-09-484-970B-142
; Sequence 142, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmut, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 142
; LENGTH: 2612
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No. 6426186 412477-1CB1
; US-09-484-970B-142

Query Match 4.2%; Score 136.2; DB 4; Length 2612;
Best Local Similarity 57.5%; Pred. No. 1.5e-24;
Matches 265; Conservative 0; Mismatches 193; Indels 3; Gaps 1;
QY 57 TAGCATGTGCTGTCCGGCTGCTGTGATCGGGGACTCCGGGTGGGCAAGACCTGCGCTG 116
Db 91 TACGACCTCAGCGGCAAGGTGATGCTTCTGGGAGACACAGCGGTGCGCAAAACATGTTTC 150
QY 117 CTGTGCGGCTTACCGACACAGAGTTCACACTCCT---CGCACATCTCCACCATCGGTGT 173
Db 151 CTGATCCCAATTCAAGACGCGGCTTCTGTCCGGAACCTTCATAGCCACCGTTCGGCATA 210
QY 174 GACTTTAAAGATGAAGACCATTAGAGGTAGACGGGATCAAAAGTGGGATACAGATCTGGGAC 233
Db 211 GACTTCAGGAACCAAGGTGGTGAAGTGGGATGCGGTGAGAGTGAAGCTGCAGATCTGGGAC 270
QY 234 ACTGAGGCGGAGAGATACAGACCATCACAAAGCAGTACTATCGGCGGGCCAGGGG 293
Db 271 ACCGCTGGGAGGAAACGGTTCGAAAGCTGCCCATGCTTATTACAGAGATGCTCGGCC 330
QY 294 ATATTTTGGTCTATGACATTAGCAGCGAGCGCTCTTACCAGCACATCATGAAGTGGGTC 353
Db 331 TTGCTTCTGCTGATGACATCACCAAAATCTCTTTGACAAACATCATCAGGCGCTGGCTC 390
QY 354 AGTGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 413

Db 391 ACTGAGATCATGAGATGCGCCAGAGGGAGCTGGTGATCATGCTGTAGCAACAAGGCG 450
QY 414 GATGAGGAGCAGAGGCGGAGGTGGGAAGAGAGCAGAGGCGAGCTGGCGAAGAGGATAT 473
Db 451 GATATGAGCAGCGAAGAGGTGATCCGTTCGGAAGACGAGAGACCTTGGCCAGGAGTAC 510
QY 474 GGCATGGAGCTTCTATGAACAAGTGCCTGCACCAACTCAA 514
Db 511 GGTGTTCCTTCTCTGGAGACCAGCGCCAAAGACTGGCATGAA 551

RESULT 7

US-08-773-423-4

; Sequence 4, Application US/08773423
; Patent No. 5869291
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; APPLICANT: Bandman, Olga
; TITLE OF INVENTION: NOVEL RAB PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/773,423
; FILING DATE: Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0183 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 847 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: Consensus

Query Match 3.7%; Score 120.6; DB 2; Length 847;
Best Local Similarity 51.5%; Pred. No. 7.3e-21;
Matches 301; Conservative 0; Mismatches 280; Indels 3; Gaps 1;

QY 22 TCCCGCCCGCTGGCCCGCCAGTCATGCGGAAGCAGTACGATGCTGTTCCGCGCTGTCG 81
Db 60 TCGCGANCAAGATGGGGAATGGAACCTGAGGAATATATACCTTGTCTTCAAGGTGGTGC 119
QY 82 TGATCGGGGACTCCGGGGTGGGCAAGACCTGCTGTGTGCGCCGCTTCACGCAACAGT 141
Db 120 TGATCGCGAATCAGGTGTGGGGAAGCAACCTACTCTCCGATTCACGCGCAATGAGT 179
QY 142 TCCATCCCTCCGCACATCTCCACCATCGGTGTTGACTTTAGATGAAGACCATAGAGTAG 201

Db 180 TCAGCCACGACACCGCCACCACCATCGGGGTGAGTTCTCCACCCCGCACTGTGATGTGG 239
QY 202 ACGGCATCAAAAGTGGGATACAGATCTGGGCACTCTGAGGCGCAGGAGAGATACACAGCCA 261
Db 240 GCACCGCTGCTGTCAAGGCTCAGATCTGGGACACAGCTGGCCTGGAGCGGTACCGAGCCA 299
QY 262 TCACAAAGCAGTACTACTGCGGGGCGCCAGGGGATATTTTGGTCTATGACATTAGCAGCG 321
Db 300 TCACCTCGGGCTACTACTGCTGCTGAGTGGGGGCCCTCTCTGGTGTGTTGACCTAACCAAGC 359
QY 322 AGCGCTCTTACCAGCACATCATGAGTGGGTCACTGACGTGGATGAGTACGCCACCAAG 381
Db 360 ACCAGACCTATGCTGTGGAGCGATGGCTGAAGGAGCTCTATGACCATGCTGAAGCCA 419
QY 382 CGCTCCAGAAAGATCCTTATTGGGAATAAGGCTGATGAGGAGCAGAAACCGAGTGGAA 441
Db 420 CGATCGTCGTCATGCTCGTGGTAAACAAAGTGAACCTCAGCCAGGCGCGGAGTGCCCA 479
QY 442 GAGAGCAAGGCGCAGCAGCTGGCGAAGGAGTATGGCATGGACTTCTATGAACAAGTGCCT 501
Db 480 CTGAGGAGGCCGGAATGCTGCTGAAACAATGAGACTGCTCTCTCTGGAGACCTCAGCCC 539
QY 502 GCACCAACCTCAACATTA--AGAGTCATTCAAGCGCTCTGACAGACTGCTGTCAGG 558
Db 540 TGGACTCTACCAATGTTGAGCTAGCCCTTGAGACTGTCTTGAAGAAATCTTTGCGAAGG 599
QY 559 CCATAGGAAGAGGCTCGGAGGCTGCGGATGGTGGCAGCAAT 602
Db 600 TGTCGAAGCAGACAGACAGCATCGGACCAATGCCATCACT 643

RESULT 8

US-09-399-913-66

; Sequence 66, Application US/09399913
; Patent No. 6361971
; GENERAL INFORMATION:
; APPLICANT: Rhodes, Kenneth
; APPLICANT: Betty, Maria
; APPLICANT: Ling, Hual-Ping
; APPLICANT: An, Wenqian
; TITLE OF INVENTION: POTASSIUM CHANNEL INTERACTORS AND USES THEREFOR
; FILE REFERENCE: MNI-070CP2
; CURRENT APPLICATION NUMBER: US/09/399,913
; CURRENT FILING DATE: 1999-09-21
; EARLIER APPLICATION NUMBER: USSN 60/110,277
; EARLIER FILING DATE: 1998-11-30
; EARLIER APPLICATION NUMBER: USSN 60/110,033
; EARLIER FILING DATE: 1998-11-25
; EARLIER APPLICATION NUMBER: USSN 60/109,333
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: USSN 09/298,731
; EARLIER FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: USSN 09/350,614
; EARLIER FILING DATE: 1999-07-03
; EARLIER APPLICATION NUMBER: USSN 09/350,874
; EARLIER FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 66
; LENGTH: 639
; TYPE: DNA
; ORGANISM: Rattus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(636)
US-09-399-913-66

Query Match 3.5%; Score 113; DB 4; Length 639;
Best Local Similarity 52.2%; Pred. No. 5.2e-19;
Matches 251; Conservative 0; Mismatches 230; Indels 0; Gaps 0;
QY 51 AAGCAGTACGATGCTGCTCCGGCTGCTGCTGATCGGGGACTCCGGGGTGGGCAAGACC 110

1 ATGGGTACGCTATCTTCAAGTACATCATCGCGACACAGGTGTTGGTAATCG 60
111 TGCCTGCTGCGGCTTCCACGACACAGAGTTCCACTCCTCGACATCTCCACCATGGT 170
61 TCGTTATGCTACAGTTTACAGACAAAGAGTTTTCAGCGGTGTCATGACCTCACAATTGGT 120
171 GTTGACTTTAAGATGAACACCATAGAGGTAGACGGCATCAAGTGGCGATACAGATCTGG 230
121 GTAGAGTTTGGTGTGCGAATGATACCAATGATGGGAAACAGATAAATCTCCATCTGG 180
231 GACACTGACGAGCGAGAGATACAGACCATCAAAAGCAGTACTATFCGCGGCGCCAG 290
181 GATACAGCAGGCGCAGAGTCTTCTGTTCTATCAACAAGTCTATATACAGAGGTGCAGCG 240
291 GGGATATTTTGGTCTATGATACATAGCAGCGCGCTTTACCAGCAGCATCATGAGTGG 350
241 GGGGCTTTTACTAGTCTATGATATACAGAGGAGACAGGTTCAACCACTTGACACCTGG 300
351 GTCAGTGACGTGGATGAGTACGACACAGAGCGGTCCAGAGATCCTTTATGGGAATAAG 410
301 TTAGAAGACGCGCGTCAAGTCCAAATCCACATGGTCTATCATGCTTATTGGAAATAA 360
411 GCTGATGAGCAGCAGAAACGCGAGGTGGAGAGAGCAAGGCGACGAGTGGCGGAGGAG 470
361 AGTGACTTAGAATCTAGGAGAGAGTGAAGAGGAGAGGTGAAGCTTTTGCAGGAGAG 420
471 TATGCGATGGACTTCTATGAACAAAGTGCCTGCACCAACCTCAACATTAAGAGATTC 530
421 CATGACTTATCTTATGTAAGAACTCTGCCAAGACTGCTTCTATGATAGAGGAGCAATT 480
531 A 531
481 A 481

RESULT 9

US-08-888-077A-28
; Sequence 28, Application US/08888077A
; Patent No. 6020143
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE AND USES THEREFOR.
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LERNER, DAVID, LITTENBERG, KRUMHOLZ & MENTLIK
; STREET: 600 SOUTH AVENUE WEST
; CITY: WESTFIELD
; STATE: NJ
; COUNTRY: USA
; ZIP: 07090-1497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/888,077A
; FILING DATE: 03-JUL-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/592,541
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: PALISI, THOMAS M
; REGISTRATION NUMBER: 36,629
; REFERENCE/DOCKET NUMBER: SCHERING 3.0-017 CIP CIP IV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 654-5000
; TELEFAX: (908) 654-7866

; INFORMATION FOR SEQ ID NO: 28:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 970 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..970
; OTHER INFORMATION: /note= "Y2H9"
US-08-888-077A-28

Query Match 3.4%; Score 110; DB 3; Length 970;

Best Local Similarity 52.9%; Pred. No. 3.6e-18;

Matches 236; Conservative 0; Mismatches 210; Indels 0; Gaps 0;

QY 55 AGTACGATGCTCTTCCGGCTGCTGCTGATCGGGGACTCCGGGTGGGCAAGACCTGCC 114
DB 73 AGTACGACTACCTCTTTAAAGTTGCTCTTATTTGGAGATTCTGGTGTGGAAAGAGTAATC 132
QY 115 TGTGTGCGCGCTTCCACGACAAACGAGTTCCACTCCTCGCACATCTCCACCATCGGTGTG 174
DB 133 TCTGTCTCGATTACTCGAATGAGTTTAATCTGGAAGCAAGAGCACCATTTGGAGTAG 192
QY 175 ACTTTAGATGAAGACCATAGAGGTAGAGGATCAAGTCCGATACAGATCTGGGACA 234
DB 193 AGTTTGAACAAGAGCATCCAGGTTGATGGAAGAAACAATAAAGGCAAGATATGGGACA 252
QY 235 CTGAGGCGAGGAGATACAGACCATCAAAAGCAGTACTATCGCGGCGCCAGGGGA 294
DB 253 CAGAGGCGAGGAGATATCGAGCTATACATCAGCATATATTCGTGGAGCTTAGTG 312
QY 295 TATTTTGGTCTATGACATTAGCAGCGGCTCTTACCAGCATCATGAAAGTGGTCA 354
DB 313 CTTTATGTTTATGACATTCCTAAACATCTCACATATGAAATCTAGAGCGATSGCTGA 372
QY 355 GTGACGTGGATGAGTACGACACAGAGCGTCCAGAGATCCTTATTGGGAATAGGCTG 414
DB 373 AAGAACTCAGAGATCATGCTGATAGTAACATTGTTATCATCTTGTGGGCATTAAGAGTG 432
QY 415 ATGAGGAGCAGAAACGCGAGGTGGGAGAGAGCAAGGCGACAGCTGGCGAAGGAGTATG 474
DB 433 ATCTACGTCTCTCAGGCGAGTTCTCTACAGATGAAGCAAGAGCTTTTCAGAAAAAGATG 492
QY 475 GCATGGACTTCTATCAACAAGTGGCC 500
DB 493 GTTTGCTATTCATTGAAACTTCGGGCC 518

RESULT 10

US-08-741-411-6
; Sequence 6, Application US/08741411
; Patent No. 6124116
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: NOVEL RAB PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/741,411
; FILING DATE: Herewith

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0139 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 820 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
LIBRARY:
CLONE: Consensus
US-08-741-411-6

Query Match 2.8%; Score 91.2; DB 3; Length 820;
Best Local Similarity 52.3%; Pred. No. 1.8e-13;
Matches 226; Conservative 0; Mismatches 203; Indels 3; Gaps 1;
QY 167 CGGTGTTGACATTTAAGATGAAGACCATAGAGTAGAGCGGCATCAAGTGGCGATACAGAT 226
DB 218 CGTGGTGGCTTCTTCAAGAGAGGTGGATGTGGTGGCCACCTCTCTGAAGCTTGAGAT 277
QY 227 GTGGGACACTGCGGCGAGGAGATACACAGACCATCACAAAGCAGTACTATCGCGGGC 286
DB 278 CTGGGACACAGCTGGCCAGGAGAGTACCACAGCGTCTGCCACCTCTACTTCAGGGGTGC 337
QY 287 CCAGGGGATATTTTGGTCTATGACATAGCAGCGGCGCTCTTACACAGCACATCAAGAA 346
DB 338 CAACGCTGCGGCTTCTGGTGTACGACATCACAGGAAGGATTCTCTCAAGGCTCAGCA 397
QY 347 GTGGTGTAGTACGCTGATGATGTA---CGCACAGAGGCGTCCAGGAAGTCCCTTATGG 403
DB 398 GTGCTGAGACCTGAGGAGGAGCTGCACCCAGGAGAGTCTCTGTGATGCTGGTGGG 457
QY 404 GAATAGGCTGTGAGGAGCAGAAACCGGAGGAGGAGAGAGGAGGAGGAGGAGGAGGAGG 463
DB 458 CAACAAGACGACCTCAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 517
QY 464 GAAGGATGTGGCATGACTTCTATGAACAAAGTGCCTGCACCAACCTCAACATTAAGA 523
DB 518 CGACAGCCAGAGTGTGTTTCATGGAACCTTCGCGCAAACTGAACCAACAGGCTCGGA 577
QY 524 GTCAATTCAGCGCTGTACAGAGCTGGTGTCTCAGCGCCCATAGGAAGGAGCTGGAAGGCT 583
DB 578 GGTGTTCAATACAGTGGCCAGAGAGTACTCCAGAGAGGAGGAGGAGGAGGAGGAGGCT 637
QY 584 CCGGATGCGTC 595
DB 638 ACGGGGATGC 649

RESULT 11
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner

STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 INMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-F15
US-08-232-463-14

Query Match 2.6%; Score 85.8; DB 1; Length 7218;
Best Local Similarity 2.3%; Pred. No. 1.1e-11;
Matches 9; Conservative 251; Mismatches 123; Indels 0; Gaps 0;
QY 804 TAGCCACTACTCCCTGCTGCTGCTGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 863
DB 1060 TTGGGATTT 1119
QY 864 TGTCCAGCCGCTCCACCTGGAGTGGTCTTCTTCCAGCTGTTTCCCGCAGCCAGCCAGCC 923
DB 1120 YY 1179
QY 924 TGCTAGACCCCGAGGATGTCGCGAAGCACTGCTCACCATCCCGCAGCCAGCCAGCAAA 983
DB 1180 YY 1239
QY 984 CAGCCAGGCTGGATCCAGGCCACTTTCAGCTGCTCTTCTCCTGTCATGCTGCTCT 1043
DB 1240 YY 1299
QY 1044 TCTGCTTTTCTCTCTTCCCGCCTCTTCTTCTGACCCCTCCCTCCGGTGGCTT 1103
DB 1300 YY 1359
QY 1104 TCGTATCAAGCTCTCAACCCCGTCCCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1163
DB 1360 YY 1419
QY 1164 TCT 1186
DB 1420 YYYYYYYYYYYYYGTACCAAA 1442

RESULT 12
US-08-741-411-4
Sequence 4, Application US/08741411

Patent No. 6124116
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Au-Young, Janice
TITLE OF INVENTION: NOVEL RAB PROTEINS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/741,411
FILING DATE: Herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0139 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 890 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
LIBRARY:
CLONE: Consensus
US-08-741-411-4

Query Match 2.5%; Score 80.6; DB 3; Length 890;
Best Local Similarity 54.0%; Pred. No. 8.4e-11;
Matches 183; Conservative 2; Mismatches 151; Indels 3; Gaps 1;
QY 30 CGCGTGGCCAGTCATGCGGAGCAGTAGATGCTGTTCCGGCTGCTGCTGCTGCGG 89
DB 36 CCGCGCCGCGAGGATGCGAGCCCGCCGACCAAGGAGCACCTGTACAAAGTTGCTGTTGGC 95
QY 90 GACTCCGGGTGGGCAAGACCTGCCCTGTGTGCGCTTCACCGACAAACAGTTTCCACTCC 149
DB 96 GACTCGGGGTGGGSAAGACCAAGATCAACAGCGTAGCTCCACCGCTCTCTCCAG 155
QY 150 TCGACATCTCCACCATCGGTGTGACTTTAAGATGAAGACCATAGAGTAGAGCGCATC 209
DB 156 CACTACCGGGCCACCATCGGGGTGGACTTCGCCCTCAAGTCTCAACTGGGACAGG 215
QY 210 A---AAGTCGGATACAGATCTGGGACACTGCGGAGCAGGAGAGATACAGACCATCA 266
DB 216 ACTCTGTGGCTGCGCTGAGCTGTGGGACATCGCGGGGAGGAGGATTTGGCAACATGACC 275
QY 267 AAGCAGTACTATCGCGGGGCCAGGAGGATTTTGTCTATGACATTTAGCAGCGAGCG 326
DB 276 CGAGTATACTACAAGGAAGCTGTGTGCTTTTGTAGTCTTTGTATATCAAGAAGTTCC 335
QY 327 TCTTACAGACATCATGAAGTGGGTGCTGAGTGGAT 365
DB 336 ACATTTGAGGCGACTCTTAAATGAAAAGTGTGATCTGGAT 374

RESULT 13
US-09-075-454-8
Sequence 8, Application US/09075454
Patent No. 6391580
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Tang, Y. Tom
APPLICANT: Lal, Preeti
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
APPLICANT: Patterson, Chandra
APPLICANT: Batra, Sajeev
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: RAS PROTEINS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Word Perfect 6.1/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/075,454
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/766,551
FILING DATE: DECEMBER 12, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Carrone, Michael C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0168-1 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1172 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: KIDNOT05
CLONE: 627565
US-09-075-454-8

Query Match 2.5%; Score 80; DB 4; Length 1172;
Best Local Similarity 54.4%; Pred. No. 1.3e-10;
Matches 161; Conservative 0; Mismatches 135; Indels 0; Gaps 0;
QY 65 GCTGTTCGGCTGCTGCTGATCGGGACTCCGGGGTGGGCAAGACCTGCTGCTGCGG 124
DB 178 GATCTCCAAGGTCATTTGTTGGTGGGACCTGTCGGTGGGGAAGACTTGCTCATTAATAG 237
QY 125 CTTACCGCAGACGAGTTCCTCTCCGACATCTCCACCATCGGTGTGACTTTAAGAT 184
DB 238 GTTCTGAAGACACCTTTTGATGAAGATTTACAAGGCCACCATTTGGAGTGGACTTCGAGAT 297
QY 185 GAAGACCATAGAGGTAGACGGCATCAAAGTGGCGGATACAGATCTGGGACACTGCAGGGCA 244
DB 298 GGAAGATTGGAGTCTGGGCAATCCCTTCAGTTTGGAGCTTGGGATACCGCTGGGCA 357
QY 245 GGAGAGATACAGACCATCAAGACGATCTATCGCGGGGCCAGGGGATATTTTGT 304
DB 358 GGAGAGGTTCAAATGCATTCATCACTACTATAGAGGAGCTCAAGCCATCATCTTGT 417

Db	658	GGAACGATTTGAGTGCTGGGCATTCCTTCAGTTTGAGGATTTGGGATACCGTGGGCA	717
Qy	245	GGAGAGATACCAGACCATCAACAAAGCAGTACTATCGCGGGCCAGGGGATATTTTTGGT	304
Db	718	GGAGAGGTTCAATGCAATTGCAATCAACCTACTATAGAGGAGCTCAAGCCATCATCATGTT	777
Qy	305	CTATGACATTAGCAGCGAGCGCTCTTACCACCATCATGAAGTGGGTTCAGTGAGC	360
Db	778	CTTCAACCTGAATGATGTGGCATCTCTGGACACATACCAAGCAGTGGCTGGCCGATG	833

Search completed: June 23, 2003, 13:39:29
Job time : 144 secs

241 GGCAGGAGAGATACCGAGACCATCACAAAGCAGTACTATCGGGGGGCCAGGGGATATTT 30 QY

RESULT 5

```

US-09-817-198A-29
; Sequence 29, Application US/09817198A
; Patent No. US20020146758A1
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001188
; CURRENT APPLICATION NUMBER: US/09/817,198A
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Homo sapien.
US-09-817-198A-29

```

RESULT 6

US-09-817-198A-30
; Sequence 30, Application US/09817198A

RESULTS

```

RESULT /
US-09-764-868-507
; Sequence 507, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: ET32
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper

```

RESULT 6

US-09-817-198A-30
; Sequence 30, Application US/09817198A

Db 393 ATGAATCAGGATCTATAGACTGATAGGCCCTATCCACAAGGGCCATGACTGGGAAAG 334
Qy 2910 GTATGGAGCAGAGGAGAAATGGATTTAGGTGCAGTACGCTCACCTCAACATTTT 2969
Db 333 GTATGGAGCAGAGGAGAAATGGATTTAGGTGCAG-TAGGCTCACCTCAACATTTT 275
Qy 2970 GGTGGCTGGGGATGCTTGGAGCCAGACTGTTAGCAGGCTCTGCTGGCTGTTTAC 3029
Db 274 GGTGGCTGGGGATGCTTGGAGCCAGACTGTTAGCAGGCTCTGCTGGCTGTTTAC 215
Qy 3030 TCGTCACCACTCTGCACCTGCTGCTTTGAGACTCCATCCAGGCCAGCGCCACCTG 3089
Db 214 TCGTCACCACTCTGCACCTGCTGCTTTGAGACTCCATCCAGGCCAGCGCCACCTG 155
Qy 3090 CTCTGAGCTCCACTATCTCCCTGTGACGGGTGAACCTCGTGTACTGTCTCGGGTCC 3149
Db 154 CTCTGAGCTCCACTATCTCCCTGTGACGGGTGAACCTCGTGTACTGTCTCGGGTCC 95
Qy 3150 ATATATGAATGTGACGAGGTTCATCTATTTTAAACACAGATGTTTACAAATAAGAT 3209
Db 94 ATATATGAATGTGACGAGGTTCATCTATTTTAAACACAGATGTTTACAAATAAGAT 35
Qy 3210 TATTTCACCAACCAAAAAAATAAAAAA 3243
Db 34 TATTTCACCAACCAAAAAAATAAAAAA 1

RESULT 12

US-09-918-995-20848
; Sequence 20848, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20848
; LENGTH: 458
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(458)
; OTHER INFORMATION: n = A,T,C or G

Query Match 13.0%; Score 424.4; DB 9; Length 458;
Best Local Similarity 99.5%; Pred. No. 1.9e-109;
Matches 425; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 2540 AGGACAGCTAACTCTAGGATGAAGCTGTGACTAGGCTGGAGTGTCTCTTGAAGATG 2599
Db 32 AGGACAGCTAACTCTAGGATGAAGCTGTGACTAGGCTGGAGTGTCTCTTGAAGATG 91
Qy 2600 GGACTCTTGGGTATCAAGACCTATGCCACATCACACTGGGGCTAGGGAAGTAGGTATG 2659
Db 92 GGACTCTTGGGTATCAAGACCTATGCCACATCACACTGGGGCTAGGGAAGTAGGTATG 151
Qy 2660 CCAGCCCTCAAGTCTGCTTTCAGCCAGGACTTGAAGAATATATTTGGCAGTGGCTCCA 2719
Db 152 CCAGCCCTCAAGTCTGCTTTCAGCCAGGACTTGAAGAATATATTTGGCAGTGGCTCCA 211
Qy 2720 ATCTGTGGCAGGATTTTCACTTTCCTTGAAGATCAGGAGGTGTCATTTGCTT 2779
Db 212 ATCTGTGGCAGGATTTTCACTTTCCTTGAAGATCAGGAGGTGTCATTTGCTT 271
Qy 2780 TCTCTCTAGCCCTCAGGAAGAGGACTATATTTGTACTGTACCCCTAGGGGTTCTGG 2839

Db 272 TCCTCTAGCCCTCAGGAAAGAGGACTATATTTGTACTGTACCTTAGGGGTTCTGG 331
Qy 2840 AAGGAAACATGAATCAGGATCTATAGACTCATAGGCCCTATCCACAAGGGCCATGA 2899
Db 332 AAGGAAACATGAATCAGGATCTATAGACTCATAGGCCCTATCCACAAGGGCCATGA 391
Qy 2900 CTGGAAAGATGTTGGAGCAGAGGAGAAATTTAGGTGTCAGCTACCTCACC 2959
Db 392 CTGGAAAGATGTTGGAGCAGAGGAGAAATTTAGGTGTCAGCTACCTCACC 451
Qy 2960 CTAAACT 2966
Db 452 CTAACT 458

RESULT 13

US-10-046-935-1930
; Sequence 1930, Application US/10046935
; Patent No. US20020156011A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.527C1
; CURRENT APPLICATION NUMBER: US/10/046,935
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2239
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1930
; LENGTH: 310
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-046-935-1930

Query Match 9.4%; Score 306.8; DB 9; Length 310;
Best Local Similarity 99.4%; Pred. No. 2.6e-76;
Matches 308; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 2824 ACCCTAGGGTCTCTGGAAGGAAACATGGAATCAGGATCTATAGACTGATAGGCCCTA 2883
Db 1 ACCCTAGGGTCTCTGGAAGGAAACATGGAATCAGGATCTATAGACTGATAGGCCCTA 60
Qy 2884 TCCACAAGGCCCATGACTGGGAAAGGATGTTGGAGCAGAGGAGAAATTTAGG 2943
Db 61 TCCACAAGGCCCATGACTGGGAAAGGATGTTGGAGCAGAGGAGAAATTTAGG 120
Qy 2944 TGCACTACGCTCACCTAACCTTTTGGTGGCTGGGGCATGCTTTCAGGCCAGACTGT 3003
Db 121 TGCACTACGCTCACCTAACCTTTTGGTGGCTGGGGCATGCTTTCAGGCCAGACTGT 180
Qy 3004 TAAGCAGGCTCTGCTGGCTGTTTACTGTCGTCACACCTCTGCACCTGCTGTTGAGACT 3063
Db 181 TAAGCAGGCTCTGCTGGCTGTTTACTGTCGTCACACCTCTGCACCTGCTGTTGAGACT 240
Qy 3064 CCATCAGCCCCCAGGACGCCACCTGCTCTGAGCCCTCCACTATCTCCCTGTGACGGGTG 3123
Db 241 CCATCAGCCCCCAGGACGCCACCTGCTCTGAGCCCTCCACTATCTCCCTGTGACGGGTG 300
Qy 3124 AACTTCGTGT 3133
Db 301 AACTTCGTGT 310

RESULT 14

US-09-878-178-1930
; Sequence 1930, Application US/09878178
; Patent No. US20020177552A1
; GENERAL INFORMATION:

```

; APPLICANT: Jlang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527
; CURRENT APPLICATION NUMBER: US/09/878,178
; NUMBER OF SEQ ID NOS: 2237
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1930
; LENGTH: 310
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-878-178-1930

Query Match
Best Local Similarity 99.4%; Score 306.8; DB 9; Length 310;
Matches 308; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2824 ACCCTAGGGTCTCTGGAGGAAACATGGAATCAGGATTTCTATAGACTGATGGCCCTA 2883
Db 1 ACCCTAGGGTCTCTGGAGGAAACATGGAATCAGGATTTCTATAGACTGATGGCCCTA 60

QY 2884 TCCACAAGGCCATGACTGGGAAAGGATATGGGAGCAGAGGAGAAATTTGGATTTAGGG 2943
Db 61 TCCACAAGGCCATGACTGGGAAAGGATATGGGAGCAGAGGAGAAATTTGGATTTAGGG 120

QY 2944 TGCAGCTACGCTCACCCTAAACATTTTGTGGCCTTGGGCGCATGTCTTGAAGGCCAGACTGT 3003
Db 121 TGCAGCTACGCTCACCCTAAACATTTTGTGGCCTTGGGCGCATGTCTTGAAGGCCAGACTGT 180

QY 3004 TAACGAGGCTCTGCTGGCCTGTTTACTCGTCACACACCTCTGCACCTGCTCTTGAGACT 3063
Db 181 TAACGAGGCTCTGCTGGCCTGTTTACTCGTCACACACCTCTGCACCTGCTCTTGAGACT 240

QY 3064 CCATCCAGCCCCAGGCACGCCACCTGCTCTGAGCCTCCACTATCTCCCTGTGACGGTG 3123
Db 241 CCATCCAGCCCCAGGCACGCCACCTGCTCTGAGCCTCCACTATCTCCCTGTGACGGTG 300

QY 3124 AACTTCGTGT 3133
Db 301 AACTTCGTGT 310
```

Search completed: June 23, 2003, 19:46:43
Job time : 445 secs

```

; APPLICANT: Jlang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527
; CURRENT APPLICATION NUMBER: US/10/146,502
; NUMBER OF SEQ ID NOS: 2241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1930
; LENGTH: 310
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-146-502-1930

Query Match
Best Local Similarity 99.4%; Score 306.8; DB 9; Length 310;
Matches 308; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2824 ACCCTAGGGTCTCTGGAGGAAACATGGAATCAGGATTTCTATAGACTGATGGCCCTA 2883
Db 1 ACCCTAGGGTCTCTGGAGGAAACATGGAATCAGGATTTCTATAGACTGATGGCCCTA 60

QY 2884 TCCACAAGGCCATGACTGGGAAAGGATATGGGAGCAGAGGAGAAATTTGGATTTAGGG 2943
Db 61 TCCACAAGGCCATGACTGGGAAAGGATATGGGAGCAGAGGAGAAATTTGGATTTAGGG 120

QY 2944 TGCAGCTACGCTCACCCTAAACATTTTGTGGCCTTGGGCGCATGTCTTGAAGGCCAGACTGT 3003
Db 121 TGCAGCTACGCTCACCCTAAACATTTTGTGGCCTTGGGCGCATGTCTTGAAGGCCAGACTGT 180

QY 3004 TAAGCAGGCTCTGCTGGCCTGTTTACTCGTCACACACCTCTGCACCTGCTCTTGAGACT 3063
Db 181 TAACGAGGCTCTGCTGGCCTGTTTACTCGTCACACACCTCTGCACCTGCTCTTGAGACT 240

QY 3064 CCATCCAGCCCCAGGCACGCCACCTGCTCTGAGCCTCCACTATCTCCCTGTGACGGTG 3123
Db 241 CCATCCAGCCCCAGGCACGCCACCTGCTCTGAGCCTCCACTATCTCCCTGTGACGGTG 300

QY 3124 AACTTCGTGT 3133
Db 301 AACTTCGTGT 310

RESULT 15
US-10-146-502-1930
; Sequence 1930, Application US/10146502
; Publication No. US20030069180A1
; GENERAL INFORMATION:
; APPLICANT: Jlang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527C2
; CURRENT APPLICATION NUMBER: US/10/146,502
; CURRENT FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1930
; LENGTH: 310
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-146-502-1930
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:23:56 ; Search time 71 Seconds
(without alignments)
397.875 Million cell updates/sec

Title: US-09-817-198b-2

Perfect score: 1105

Sequence: 1 MAKQYDLVRLLLIGDSGVG.....LEEEKGPGPANSSKTCWC 212

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1105	100.0	401	AAU17136	Novel signal trans
2	1092	98.8	218	AA341604	Human ORFX ORF1368
3	832	75.3	188	AAU17555	Novel signal trans
4	546.5	49.5	204	AB870670	Drosophila melanog
5	545	49.3	213	AB811916	Human rab8 homolog
6	540	48.9	221	AB411333	Human ovarian anti
7	532	48.1	246	AA358196	Lung cancer associ
8	530.5	48.0	200	AA191165	Amino acid sequenc
9	530.5	48.0	200	AA190979	Human Rab10 protei
10	530.5	48.0	200	AA895340	Human protein sequ

11	530.5	48.0	218	21	AA856993	Human prostate can
12	529.5	47.9	218	22	ABG07266	Novel human diagno
13	528.5	47.8	207	22	AB871647	Drosophila melanog
14	528.5	47.8	207	22	AA667154	Amino acid sequenc
15	528.5	47.8	207	22	AA92628	Human protein sequ
16	526	47.6	201	21	AA809982	Canine Rab10 prote
17	523	47.3	199	21	AA809980	Human Rab10 protei
18	523	47.3	199	21	AA809981	Human Rab10 protei
19	519	47.0	216	21	AA808688	Arabidopsis thalia
20	519	47.0	216	21	AA853945	Arabidopsis thalia
21	519	47.0	253	21	AA853944	Arabidopsis thalia
22	519	47.0	254	21	AA808687	Arabidopsis thalia
23	513.5	46.5	215	21	AA835215	Zea mays protein f
24	510	46.2	218	21	AA819220	Arabidopsis thalia
25	507	45.9	216	21	AA808006	Arabidopsis thalia
26	506.5	45.8	209	22	AB823365	Novel human diagno
27	501	45.3	224	21	AA847826	Arabidopsis thalia
28	501	45.3	234	21	AA847825	Arabidopsis thalia
29	501	45.3	335	22	AB823366	Novel human diagno
30	478.5	43.3	201	20	AA800919	Human Rab protein,
31	478.5	43.3	201	22	AA828024	Novel human secret
32	478.5	43.3	221	22	AA813525	Human polypeptide
33	478.5	43.3	224	21	AA858758	Breast and ovarian
34	476	43.1	203	21	AA830499	Arabidopsis thalia
35	476	43.1	258	21	AA830498	Arabidopsis thalia
36	468	42.4	202	21	AA807763	Arabidopsis thalia
37	467	42.3	205	21	AA834843	Arabidopsis thalia
38	467	42.3	205	21	AA834844	Gene 44 human secr
39	466	42.2	202	21	AA810858	Human secreted pro
40	462	41.8	205	22	AB859808	Arabidopsis thalia
41	457	41.4	218	21	AA830710	Drosophila melanog
42	457	41.4	221	21	AA830709	Arabidopsis thalia
43	457	41.4	254	21	AA845323	Arabidopsis thalia
44	456.5	41.3	227	21	AA819163	Amino acid sequenc
45	456.5	41.3	227	22	AA839600	Human polypeptide

ALIGNMENTS

RESULT 1

AAU17136

ID AAU17136 standard; Protein; 401 AA.

AC AAU17136;

XX XX

DT 07-NOV-2001 (first entry)

XX XX

DE Novel signal transduction pathway protein, Seq ID 701.

XX XX

KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;

KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;

KW immune system disorder; rheumatoid arthritis; inflammatory condition;

KW organ transplant rejection; infection; hepatitis C; blood disorder;

KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;

KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;

KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;

KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;

KW reproductive system; gastrointestinal; liver disorder; AIDS;

KW acquired immune deficiency syndrome.

OS Homo sapiens.

XX XX

XX WO200154733-A1.

PN PN

XX 02-AUG-2001.

PD PD

XX 17-JAN-2001; 2001WO-US01312.

PF PF

XX 31-JAN-2000; 2000US-0179065.

PR PR

XX 04-FEB-2000; 2000US-0180628.

PR PR

XX 24-FEB-2000; 2000US-0184664.

PR PR

XX 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0233397.
PR 14-SEP-2000; 2000US-0233398.
PR 14-SEP-2000; 2000US-0233399.
PR 14-SEP-2000; 2000US-0233400.
PR 14-SEP-2000; 2000US-0233401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234224.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235483.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 01-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 06-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-465460/50.
XX N-PSDB; AAS27053.
XX Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders
XX Claim 1; SEQ ID No 701; 880pp; English.
XX

CC The invention relates to novel isolated polypeptides (I), and
 CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
 CC diagnosing, preventing and treating diseases including immune system
 CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
 CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
 CC transplant rejections and graft versus host disease, infectious diseases
 CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
 CC other blood-related disorders (sickle cell anaemia), myeloproliferative
 CC disorders, primary haematopoietic disorders, hyperproliferative
 CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
 CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
 CC disorders (e.g. glomerulonephritis), cardiovascular disorders
 CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
 CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
 CC Addison's disease), reproductive system disorders, gastrointestinal
 CC disorder (inflammatory disorders), liver disorders (cirrhosis),
 CC as stimulators of B-cell responsiveness to pathogens, activators of
 CC T-cells, to induce higher affinity antibodies, and as a means to induce
 CC tumour proliferation in pathologies e.g. acquired immune deficiency
 CC syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction
 CC pathway protein, amino acid sequences of the invention.
 XX

Query Match 100.0%; Score 1105; DB 22; Length 401;
 Best Local Similarity 100.0%; Pred. No. 2.1e-108;
 Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSTIGVDFKMTIEVDGKVRQ 60
 DB 27 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSTIGVDFKMTIEVDGKVRQ 86
 QY 61 IWDTAGERYQTIKQYRRAQGIPLVYDISSERSYQHIMKWSVDVEYAPGKILIG 120
 DB 87 IWDTAGERYQTIKQYRRAQGIPLVYDISSERSYQHIMKWSVDVEYAPGKILIG 146
 QY 121 NKADEEOKRQVGRGQQLAKEYGMDVFETSACTNLNLIKESFTRTELVLQAHKLEGL 180
 DB 147 NKADEEOKRQVGRGQQLAKEYGMDVFETSACTNLNLIKESFTRTELVLQAHKLEGL 206
 QY 181 RMASNELALAELEEEGKPEGPANSSKTCWC 212
 DB 207 RMASNELALAELEEEGKPEGPANSSKTCWC 238

RESULT 2
 ID AAB41604 standard; Protein; 218 AA.
 XX AC AAB41604;
 DT 08-FEB-2001 (first entry)
 XX Human OREF1368 polypeptide sequence SEQ ID NO:2736.
 DE Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antihypertensive;
 KW antinaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; erythematoidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive.
 OS Homo sapiens.
 XX OS
 XX WO200058473-A2.
 PN

XX 05-OCT-2000.
 PD 31-MAR-2000; 2000WO-US08621.
 XX 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX (CURA-) CURAGEN CORP.
 XX Shimkets RA, Leach M;
 XX WPI; 2000-602362/57.
 DR N-PSDB; AAC75813.
 XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX Claim 11; Page 1979-1980; 5507pp; English.
 XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antifungal; antirheumatic;
 CC antihypertensive; antinaemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hyperextension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX SQ Sequence 218 AA;
 Query Match 98.8%; Score 1092; DB 21; Length 218;
 Best Local Similarity 97.2%; Pred. No. 2e-107;
 Matches 212; Conservative 0; Mismatches 0; Indels 6; Gaps 1;
 QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSTIGVDFKMTIEVDGKVRQ 60
 DB 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSTIGVDFKMTIEVDGKVRQ 60
 QY 61 IWDTAGERYQTIKQYRRAQGIPLVYDISSERSYQHIMKWSVDVEYAPGKILIG 120
 DB 61 IWDTAGERYQTIKQYRRAQGIPLVYDISSERSYQHIMKWSVDVEYAPGKILIG 120
 QY 121 NKADEEOKRQVGRGQQLAKEYGMDVFETSACTNLNLIKESFTRTELVLQAHK 174
 DB 121 NKADEEOKRQVGRGQQLAKEYGMDVFETSACTNLNLIKESFTRTELVLQAHK 180
 QY 175 KELEGLMRASNELALAELEEEGKPEGPANSSKTCWC 212
 DB 181 KELEGLMRASNELALAELEEEGKPEGPANSSKTCWC 218
 RESULT 3
 ID AAU17555 standard; Protein; 188 AA.
 XX AAU17555;
 AC AAU17555;

XX 07-NOV-2001 (first entry)
XX Novel signal transduction pathway protein, Seq ID 1120.
DE
XX
XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antinflammatory; anti-HIV; antibacterial; antinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS;
KW acquired immune deficiency syndrome.
OS
XX Homo sapiens.
XX
XX WO200154733-A1.
PN
XX
XX 02-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US01312.
PF
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216547.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226581.
PR 22-AUG-2000; 2000US-0226586.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246529.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.

XX 11-JAN-2002 (first entry)
 XX Human rab8 homologue, SEQ ID NO:2286.
 DE Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW cytotaxtic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytotaxtic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnery; antiulcer.
 XX Homo sapiens.
 OS
 PN WO200157188-A2.
 XX
 XX 09-AUG-2001.
 XX
 XX 05-FEB-2001; 2001WO-US03800.
 XX
 XX 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 XX
 XX (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Drmanac RT;
 XX WPI; 2001-457740/49.
 DR N-PSDB; ABA09160.
 XX
 XX Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer -
 XX
 PS Claim 20; Page 276; 1963pp; English.
 XX
 CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g. myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to

CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a novel human
 CC polypeptide of the invention.
 XX
 SQ Sequence 213 AA;
 Query Match 49.3%; Score 545; DB 22; Length 213;
 Best Local Similarity 52.9%; Pred. No. 2.1e-49;
 Matches 99; Conservative 45; Mismatches 4; Indels 0; Gaps 0;
 QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFSSHSHTIGVDKMKTEYDGIKVIQ 60
 DB 7 MAKTYDYLKLLIGDSGVGKTCVLFSEDAFNSTFTIGIDFKTIFELDKRKLQ 66
 QY 61 IWDYAGQERYQTIKQYRRAGQIFLVYDIDSSERYQHIMKWSVDYEAPEGVKILIG 120
 DB 67 IWDYAGQERYQTIKQYRRAGQIFLVYDIDSSERYQHIMKWSVDYEAPEGVKILIG 126
 QY 121 NKADEEQKQVREGQOOLAKKEYGMDYETACTNLNKESTRLTELVLQAHKKEGL 180
 DB 127 NKCDVNDKQVSKERGEKLDYGIKFMETSAKINVENAFFTLARDIKAKMDKKEG 186
 QY 181 RMRASNE 187
 DB 187 SPQGSNQ 193
 RESULT 6
 ID ABP41333 standard; Protein: 221 AA.
 AC ABP41333;
 DT 23-AUG-2002 (first entry)
 XX Human ovarian antigen HCGMA67, SEQ ID NO:2465.
 KW Human; ovarian antigen; ovary; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive.
 XX Homo sapiens.
 OS
 XX WO200200677-A1.
 PD 03-JAN-2002.
 XX
 XX 07-JUN-2001; 2001WO-US18569.
 XX
 XX 07-JUN-2000; 2000US-209467P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Birse CE, Rosen CA;
 XX
 DR WPI; 2002-147878/19.
 DR N-PSDB; ABO54410.
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -

XX PS Claim 11; SEQ ID No 2465; 2922pp; English.

XX CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
XX CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
XX CC encompasses polypeptides 90% identical and polynucleotides 95% identical
XX CC to the sequences of the invention. The invention additionally relates to
XX CC recombinant vectors and host cells comprising human ovarian antigen
XX CC polynucleotides, antibodies against human ovarian antigens, and the use
XX CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
XX CC treating, prognosing or preventing various ovarian and/or breast-related
XX CC disorders. Such conditions include ovarian cancer and breast cancer, and
XX CC metastatic tumours of ovarian or breast origin, reproductive system
XX CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
XX CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
XX CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
XX CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
XX CC vaginitis), immune disorders (e.g., congenital and acquired
XX CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
XX CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
XX CC respiratory disorders, neurological disorders, gastrointestinal disorders
XX CC and urinary system disorders. Ovarian antigen polypeptides and
XX CC polynucleotides may also be used in screening for compounds which
XX CC modulate ovarian antigen expression or activity. The polynucleotides may
XX CC further be used for gene therapy, chromosome mapping, in the
XX CC identification of individuals and in forensic analysis, and the
XX CC polypeptides may be used as food additives or to prepare antibodies
XX CC useful in disease diagnosis, drug targeting and phenotyping. The present
XX CC sequence represents a human ovarian antigen of the invention.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 221 AA;

Query Match 48.9%; Score 540; DB 23; Length 221;
Best Local Similarity 52.7%; Pred. No. 7.5e-49;
Matches 98; Conservative 45; Mismatches 43; Indels 0; Gaps 0;

QY 2 AKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHSTIGVDFKMTIEVDGKVRQI 61
DB 16 AKTYDYLKLLIGDSGVGKTCVLFSEDAFNSTFSTIGIDFKINTIELDGKRKLQI 75

QY 62 WDTAGQERYOTITKQYVRRAGIFLVYDTSSESYQHIMKWSVDVEYAPGVQKILGN 121
DB 76 WDTAGQERFTITTYIRGAMGIMLYDITNEKSFDRNIRNIRNIEHSAADVVEKMTIGN 135

QY 122 KADEEQKRVQREGQGOQLAKKEYGMDVETSACTNLNLIKESFTRLTETVLQAHKRELRL 181
DB 136 KCDVNDKRVQSKERGEKALDYGIKFMETSAKANINVENAFFTLARDIKAMDKKLEGNS 195

QY 182 MRASNE 187
DB 196 PGGSNQ 201

RESULT 7
AAB58196
ID AAB58196 standard; Protein; 246 AA.

XX AC AAB58196;
XX DE 14-MAR-2001 (first entry)
XX DE Lung cancer associated polypeptide sequence SEQ ID 534.

XX KW Human; lung cancer associated protein; neuroprotective; cytostatic;
XX KW cardioactive; immunomodulatory; muscular active; vulnerary;
XX KW gastrointestinal; nephrotropic; antinefective; gynecological;
XX KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
XX KW proliferative disorder; wound healing; infectious disease.
XX OS Homo sapiens.

XX PN WO200055180-A2.
XX PD 21-SEP-2000.
XX PF 08-MAR-2000; 2000WO-US05918.
XX PR 12-MAR-1999; 99US-0124270.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PA (ROSE/) ROSEN C A.
XX PI Ruben SM;
XX DR WPI; 2000-587514/55.
XX DR N-PSDB; AAF18072.
XX PT Lung cancer associated gene sequences, referred to as lung cancer
XX PT antigens, useful for treatment, prevention, and diagnosis of disorders
XX PS such as lung cancer -
XX PS Claim 11; Page 1027-1028; 1425pp; English.

XX CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
XX CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
XX CC associated proteins and polynucleotide sequences, their agonists, and
XX CC antagonists may have neuroprotective; cytostatic; cardioactive;
XX CC immunomodulatory; muscular active general; vulnerary; gastrointestinal
XX CC general; nephrotropic; antinefective; gynecological; or antibacterial
XX CC activity. The invention also includes antibodies specific for the
XX CC protein or polynucleotide sequences. The lung cancer associated
XX CC polynucleotide sequences may be used for detection of lung cancer,
XX CC chromosome identification, as chromosome markers, and for numerous other
XX CC diagnostic or research purposes. The proteins may be used to treat
XX CC disorders such as neural, immune, muscular, reproductive,
XX CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
XX CC disorders. The proteins may also be used in the treatment of wounds and
XX CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
XX CC peptide AAB58549 are used in the course of the invention for the
XX CC identification and characterisation of the polynucleotide and protein
XX CC sequences.

XX SQ Sequence 246 AA;

Query Match 48.1%; Score 532; DB 21; Length 246;
Best Local Similarity 47.6%; Pred. No. 6.2e-48;
Matches 100; Conservative 49; Mismatches 43; Indels 18; Gaps 3;

QY 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHSTIGVDFKMTIEVDGKVRQI 60
DB 44 MAKAYDHLKLLIGDSGVGKTCLLIRFAEDNFNTYISTIGIDFKINTVDIEGKIKLQ 103

QY 61 IWDTAGOERYOTITKQYVRRAGIFLVYDTSSESYQHIMKWSVDVEYAPGVQKILIG 120
DB 104 WDTAGQERFTITTYIRGAMGIMLYDITDEKSFENIQNMKSIXENASAGVERLLIG 163

QY 121 NKADEQKRVQREGQGOQLAKKEYGMDVETSACTNLNLIKESFTRLTETVLQAHKRELGL 180
DB 164 NKCDMEAKRVQKEQADKLAREHGIRFETSAKSMNVDEAFSSILDIL-----LKSQ 217

QY 181 RMASNELALAELEEEGKPEGPANSSKTC 210
DB 218 GRRSGN-----GNKP--PSTDLKTC 235

RESULT 8
AAB19165
ID AAB19165 standard; Protein; 200 AA.
XX AC AAB19165;
XX DT 19-FEB-2001 (first entry)
XX DE

DE Amino acid sequence of human RAB10.
 KW RAB protein; GTPase; GTP binding; gene therapy; cancer.
 XX Homo sapiens.
 OS
 XX WO2000058464-A2.
 PN
 XX 05-OCT-2000.
 PD
 XX 13-MAR-2000; 2000WO-US06330.
 PF
 XX 25-MAR-1999; 99US-0126083.
 PR
 XX (AXYS-) AXYS PHARM INC.
 PA
 XX Allen M, Abel K, McIntosh B, Vega R, Rutter M, Buckler AJ;
 PI WPI; 2000-647233/62.
 DR N-PSDB; AAA96887.
 DR
 XX Novel isolated nucleic acid encoding a mammalian RAB protein useful for
 PT identifying homologous or related genes, in producing composition that
 PT modulates expression or function of RAB for cancer therapy
 XX
 PS Claim 2; Page 39; 58pp; English.
 XX
 CC The present sequence represents a mammalian RAB protein. RAB proteins
 CC constitute the largest family of small GTPases, with over 40 identified
 CC isoforms. RAB proteins contain four highly conserved peptide sequences
 CC involved in GTP binding and hydrolysis. Compositions comprising RAB
 CC nucleic acid are useful for identifying homologous or related genes,
 CC in producing compositions that modulate the expression or function of
 CC RAB, for gene therapy, mapping functional regions of the protein and
 CC in studying associated physiological pathways. In addition, modulation
 CC of the gene activity in vivo is used for prophylactic and therapeutic
 CC purposes, such as treatment of cancer, and identification of cell type
 CC based on expression. The DNA may also be used to identify expression of
 CC the gene in a biological specimen.
 XX
 SQ Sequence 200 AA;
 Query Match 48.0%; Score 530.5; DB 21; Length 200;
 Best Local Similarity 55.8%; Pred. No. 6.6e-48;
 Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;
 QY 1 MAKO-YDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSHSTIGVDVFKMTIEVDGIKVI 59
 DB 1 MAKTYDLLEKLLIGDSGVGKTCVLFPSDDAFTTFTISIGIDFKIKTVLQGGKIKL 60
 QY 60 QIWDTAGQERTQITKQYRRAQGIPLYDISSERSYOHIMKWSVDVDEYAPGVQKILI 119
 DB 61 QIWDTAGQERTHTTTSYRGAMGIMLYDITNGKSFENISKWLNIHANEDVERMLL 120
 QY 120 GNKADEOKRGVREGOQOLAKYGMDFVETSACTNLNKKESFTLTELVLQ 171
 DB 121 GNKCDMDKRVVPKGEQIAREHGIRFETSANINIERAKFLTAEIDLIR 172
 RESULT 9
 AAB09979
 ID AAB09979 standard; Protein; 200 AA.
 XX
 AC AAB09979;
 XX
 DT 19-OCT-2000 (first entry)
 XX
 DE Human RAB10 protein.
 XX
 KW Human; RAB10; Yap/Rab family.
 XX
 OS Homo sapiens.

PN CN1249345-A.
 XX
 PD 05-APR-2000.
 XX
 PF 28-SEP-1998; 98CN-0121911.
 XX
 PR 28-SEP-1998; 98CN-0121911.
 XX
 PA (UYFU-) UNIV FUDAN.
 XX
 PI Yu L, Tu Q, Gao J;
 XX WPI; 2000-400723/35.
 DR N-PSDB; AAA40104.
 DR
 XX Preparation of human gene coding sequence, its encoded polypeptide -
 PT Claim 2; Page 18; 23pp; Chinese.
 PS
 XX This invention describes a novel human Rab10 cDNA sequence. The protein
 CC coded by said sequence belongs to Yap/Rab family and is the homolog of
 CC mouse Rab10. The present invention also relates to the polypeptide coded
 CC by said nucleotide sequence and the application and preparing process
 CC of said polynucleotide and said polypeptide. This sequence represents the
 CC human Rab10 protein described in the method of the invention.
 XX
 SQ Sequence 200 AA;
 Query Match 48.0%; Score 530.5; DB 21; Length 200;
 Best Local Similarity 55.8%; Pred. No. 6.6e-48;
 Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;
 QY 1 MAKO-YDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSHSTIGVDVFKMTIEVDGIKVI 59
 DB 1 MAKTYDLLEKLLIGDSGVGKTCVLFPSDDAFTTFTISIGIDFKIKTVLQGGKIKL 60
 QY 60 QIWDTAGQERTQITKQYRRAQGIPLYDISSERSYOHIMKWSVDVDEYAPGVQKILI 119
 DB 61 QIWDTAGQERTHTTTSYRGAMGIMLYDITNGKSFENISKWLNIHANEDVERMLL 120
 QY 120 GNKADEOKRGVREGOQOLAKYGMDFVETSACTNLNKKESFTLTELVLQ 171
 DB 121 GNKCDMDKRVVPKGEQIAREHGIRFETSANINIERAKFLTAEIDLIR 172
 RESULT 10
 AAB95340
 ID AAB95340 standard; Protein; 200 AA.
 XX
 AC AAB95340;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:17619.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.

PI Ota T, Isoqai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI; 2001-318749/34.
 XX
 XX
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 XX
 PS Claim 8; SEQ ID 17619; 2537pp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
 CC AAH93893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 200 AA;

Query Match 48.0%; Score 530.5; DB 22; Length 200;
 Best Local Similarity 55.8%; Pred. No. 6.6e-48;
 Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;
 QY 1 MAQKQ-YDVLFRLLIGDSGVGKTCVLFRTDNEFHSHSHISTIGVDFKMKITVEDGKIKVRI 59
 DB 1 MAKTYDLLFKLLIGDSGVGKTCVLFRTDNEFHSHSHISTIGVDFKMKITVEDGKIKVRI 60
 QY 60 QIWDTAGQERYQITKQYRRAGQIFLVYDISSERSYQHIMKWSVDVEYAPGVQKILI 119
 DB 61 QIWDTAGQERFHTTTSYRGANGIMLVYDITNGKSFENISKWLNRIDEHANEDVERMLL 120
 QY 120 GNKADEQKRVGREGQQQLAKYGMDFYETSACTNLNIKESFTRTLVELVQ 171
 DB 121 GNKCDMDKRVVPKRGQIAREHGIRFFETSANINIEKAFITLAEDILR 172

RESULT 11
 AAB56993
 ID AAB56993 standard; Protein; 218 AA.
 XX
 AC AAB56993;
 XX
 DT 13-MAR-2001 (first entry)
 XX
 DE Human prostate cancer antigen protein sequence SEQ ID NO:1571.
 XX
 KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
 KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
 KW vulnary; gastrointestinal; nephrotropic; antineoplastic; gynaecological;
 KW antibacterial; gene therapy; neural; immune; reproductive; renal;
 KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
 KW wound; infectious disease.
 XX
 OS Homo sapiens.
 XX
 XX

PN WO200055174-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05988.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSE/) ROSEN C A.
 XX
 PI Rosen CA, Ruben SM;
 XX
 XX WPI; 2000-587513/55.
 DR N-PSDB; AAF16196.
 DR
 XX
 PT Prostate cancer associated gene sequences, referred to as prostate
 PT cancer antigens, useful for treatment, prevention, and diagnosis of
 PT disorders such as prostate cancer -
 XX
 PS Claim 11; Page 2014-2015; 2338pp; English.
 XX
 CC AAF15566 to AAF16505 encode the human prostate cancer associated
 CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
 CC The prostate cancer antigens can have neuroprotective, cytostatic,
 CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
 CC nephrotropic, antineoplastic, gynaecological and antibacterial activities,
 CC and can be used in gene therapy. The prostate cancer antigen
 CC polynucleotides may be used for detection of prostate cancer
 CC identification, as chromosome markers, and for numerous other diagnostic
 CC or research purposes. The prostate cancer antigens may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
 CC AAB57303 represent sequences used in the exemplification of the present
 CC invention.
 XX
 SQ Sequence 218 AA;

Query Match 48.0%; Score 530.5; DB 21; Length 218;
 Best Local Similarity 55.8%; Pred. No. 7.5e-48;
 Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;
 QY 1 MAQKQ-YDVLFRLLIGDSGVGKTCVLFRTDNEFHSHSHISTIGVDFKMKITVEDGKIKVRI 59
 DB 19 MAKTYDLLFKLLIGDSGVGKTCVLFRTDNEFHSHSHISTIGVDFKMKITVEDGKIKVRI 78
 QY 60 QIWDTAGQERYQITKQYRRAGQIFLVYDISSERSYQHIMKWSVDVEYAPGVQKILI 119
 DB 79 QIWDTAGQERFHTTTSYRGANGIMLVYDITNGKSFENISKWLNRIDEHANEDVERMLL 138
 QY 120 GNKADEQKRVGREGQQQLAKYGMDFYETSACTNLNIKESFTRTLVELVQ 171
 DB 139 GNKCDMDKRVVPKRGQIAREHGIRFFETSANINIEKAFITLAEDILR 190

RESULT 12
 AAB56993
 ID AAB56993 standard; Protein; 218 AA.
 XX
 AC AAB56993;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #7257.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 XX WO200175067-A2.
 XX


```

PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Drmanac RT, Liu C, Tang YT;
PI
XX WPI; 2001-639362/73.
DR N-PSDB; AAS71453.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
XX Claim 20; SEQ ID NO 37625; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 218 AA;
SQ
Query Match 47.9%; Score 529.5; DB 22; Length 218;
Best Local Similarity 55.8%; Pred. No. 9.6e-48;
Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;
QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDFKMTIEVDGIKVI 59
DB 19 MAKTYDYLFKLLIGDSGVGKTCVLFPSDDAENFTFSTIGIDFKIKTVELQGGKIKL 78
QY 60 QIWDTAGOERYOTITKQYRRAGQIFLVYDYSERSYOHIMKWSVDVEYAPGVOKILI 119
DB 79 QIWDTAGOERHTTITSYRGAMGLVYDITNGKSPENISKWLINDEHANEDVERMLL 138
QY 120 GNKADEQKRVGREGQOOLAKKEYGMDFYETSACTNLNIKESFTRLTELVLQ 171
DB 139 GNKCDMDKRVVPKRGQIAREHGIRFETTSKANINIEKAFLLTAEILR 190
RESULT 13
ABB71647
ID ABB71647 standard; Protein: 207 AA.
XX
XX ABB71647;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 41733.
DE
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW

```

```

KW pharmaceutical.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US09231.
PF
XX
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI; 2001-656860/75.
DR N-PSDB; ABL15750.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions
XX
XX Disclosure; SEQ ID NO 41733; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 207 AA;
SQ
Query Match 47.8%; Score 528.5; DB 22; Length 207;
Best Local Similarity 53.8%; Pred. No. 1.1e-47;
Matches 100; Conservative 41; Mismatches 40; Indels 5; Gaps 1;
QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDFKMTIEVDGIKVI 60
DB 1 MAKTYDYLFKLLIGDSGVGKTCILFRSEDAFNTFTSTIGIDFKIKTIELDNKKIKLQ 60
QY 61 IWDTAGOERYOTITKQYRRAGQIFLVYDYSERSYOHIMKWSVDVEYAPGVOKILIG 120
DB 61 IWDTAGOERFRTITAYYRGAMGLVYDITQKSFENIKWIRNIEENASADVKEKMLLG 120
QY 121 NKABDEQKRVGREGQOOLAKKEYGMDFYETSACTNLNIKESFTRLTELVLQAHKKEGL 180
DB 121 NKCELTDRQVSKERGEQLAIEYGKFKMETSAKASINVERAP-----LTASDIKAKTEK 175
QY 181 RMRASN 186
DB 176 RMEANN 181
RESULT 14
AAG67154
ID AAG67154 standard; Protein: 207 AA.
XX
XX AAG67154;
AC
XX
XX 13-NOV-2001 (first entry)
DT
XX
XX Amino acid sequence of human 27423 G-protein.
DE
XX
XX Human; G-protein; 32705; 23224; 27423; 32700; 32712; lung disorder;
KW congenital anomaly; pulmonary congestion; oedema; haemorrhage;

```

KW adult respiratory distress syndrome; Goodpasture's syndrome;
 KW chronic obstructive pulmonary disease; asthma; pulmonary hypertension;
 KW liver disorder; hepatic injury; jaundice; cholestasis; viral hepatitis;
 KW cirrhosis; Wilson's disease; autoimmune hepatitis; hepatic failure;
 KW brain disorder; hypoxia; cerebral ischemia; intracranial haemorrhage;
 KW acute meningitis; Parkinson's disease; Alzheimer's disease; glioma;
 KW chronic bacterial meningencephalitis; multiple sclerosis;
 KW ankyrotic lateral sclerosis; stroke; Huntington's disease.
 XX
 OS Homo sapiens.
 XX
 PN W0200164887-A2.
 XX
 PD 07-SEP-2001.
 XX
 XX 27-FEB-2001; 2001WO-US06292.
 XX
 XX 29-FEB-2000; 2000US-0185606.
 XX
 XX (MILL-) MILLENNIUM PHARM INC.
 PA
 PI Meyers RA;
 XX
 XX WPI: 2001-550182/61.
 DR N-PSDB; AAG87154.
 DR
 XX Novel human small G-protein polypeptides and polynucleotides for
 PT treating lung disorders, liver disorders and brain disorders -
 PT
 XX Claim 8; Fig 15; 151pp; English.

XX The present sequence represents a human G-protein. The specification
 CC describes 32705, 23224, 27423, 32700 or 32712 small G-proteins. The
 CC G-protein polypeptides and polynucleotides are useful as a target for
 CC diagnosis and treatment of G-protein mediated or related disorders,
 CC and for identifying agonists and antagonists for diagnosis and
 CC treatment. They are useful for treating disorders of lung (e.g.
 CC congenital anomalies, pulmonary congestion, oedema, adult respiratory
 CC distress syndrome, haemorrhage, chronic obstructive pulmonary disease,
 CC asthma, Goodpasture's syndrome and pulmonary hypertension), liver
 CC (e.g. hepatic injury, jaundice, cholestasis, viral hepatitis, cirrhosis,
 CC Wilson's disease, autoimmune hepatitis and hepatic failure), and
 CC brain (e.g. hypoxia, cerebral ischemia, intracranial haemorrhage, acute
 CC meningitis, Parkinson's disease, Alzheimer's disease, gliomas, chronic
 CC bacterial meningencephalitis, multiple sclerosis, amyotrophic lateral
 CC sclerosis, stroke and Huntington's disease).

SQ Sequence 207 AA;
 Query Match 47.88; Score 528.5; DB 22; Length 207;
 Best Local Similarity 54.08; Pred. No. 1.1e-47;
 Matches 95; Conservative 47; Mismatches 33; Indels 1; Gaps 1;
 QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHISTIGVDKMKTEVDGKVKRIQ 60
 DB 1 MAKTYDLFLKLLIGDSGVGKTCCLFRFSDAFTNTFTSTIGDKRTIELDGKKIKLQ 60
 QY 61 IWDTAGQERYQTITKQYRRAQGIYFLYDISESSYQHIMKWSVDVEYAPGKQILIG 120
 DB 61 IWDTAGQERYQTITAYYRGAMGIMLYDITNEKSPDNKWNIRNIEHASSDVERMILG 120
 QY 121 NKADEOKRQVREGQGOOLAKEYGMDFEYTSACNLNIKES-FTRLTELVLQAHRK 175
 DB 121 NKCDMNDKRVSKERGEKLAIDYGIKFLTSKSSANVEAFETFLARDIMTKLNK 176

RESULT 15
 AAB92628
 ID AAB92628 standard; Protein; 207 AA.

XX
 AC AAB92628;
 XX
 XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:10930.
 DE
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 KW
 XX Homo sapiens.
 OS
 XX EP1074617-A2.
 PN
 XX 07-FEB-2001.
 XX
 XX 28-JUL-2000; 2000EP-0116126.
 PF
 XX 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 XX (HELI-) HELIX RES INST.
 PA
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI: 2001-318749/34.
 DR
 XX Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 8; SEQ ID 10930; 2537pp + CD ROM; English.
 XX The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

SQ Sequence 207 AA;

Query Match 47.88; Score 528.5; DB 22; Length 207;
 Best Local Similarity 54.08; Pred. No. 1.1e-47;
 Matches 95; Conservative 47; Mismatches 33; Indels 1; Gaps 1;
 QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHISTIGVDKMKTEVDGKVKRIQ 60
 DB 1 MAKTYDLFLKLLIGDSGVGKTCCLFRFSDAFTNTFTSTIGDKRTIELDGKKIKLQ 60
 QY 61 IWDTAGQERYQTITKQYRRAQGIYFLYDISESSYQHIMKWSVDVEYAPGKQILIG 120
 DB 61 IWDTAGQERYQTITAYYRGAMGIMLYDITNEKSPDNKWNIRNIEHASSDVERMILG 120
 QY 121 NKADEOKRQVREGQGOOLAKEYGMDFEYTSACNLNIKES-FTRLTELVLQAHRK 175
 DB 121 NKCDMNDKRVSKERGEKLAIDYGIKFLTSKSSANVEAFETFLARDIMTKLNK 176

Search completed: June 18, 2003, 15:33:30
Job time : 72 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:30:31 ; Search time 40 Seconds
(without alignments)
509.512 Million cell updates/sec

Title: US-09-817-198B-2
Perfect score: 1105
Sequence: 1 MAKQYDVLFRLLIGDSGVG.....LEEEKGPEGPANSSKTCWC 212

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1077	97.5	212	2 F42148	GTP-binding protei
2	559	50.6	224	2 T33855	hypothetical protei
3	545	49.3	207	2 B49647	GTP-binding protei
4	545	49.3	207	2 B36364	GTP-binding protei
5	532	48.1	203	2 A49647	GTP-binding protei
6	531.5	48.1	200	2 A38625	GTP-binding protei
7	530.5	48.0	200	2 D36364	GTP-binding protei
8	530	48.0	209	2 B38625	GTP-binding protei
9	526	47.6	206	2 I78851	GTP-binding protei
10	522.5	47.3	200	2 B42148	GTP-binding protei
11	519	47.0	216	2 T45901	GTPase ALRAB8 - Ar
12	514	46.5	201	2 T38971	hypothetical protei
13	512	46.3	216	2 S33900	GTP-binding protei
14	512	46.3	216	2 T48378	GTP-binding protei
15	510	46.2	200	2 S12790	GTP-binding protei
16	509	46.1	215	2 T14565	GTP-binding protei
17	508.5	46.0	215	2 S7478	GTP-binding protei
18	508	46.0	216	2 J30640	GTP-binding protei
19	507.5	45.9	222	2 T14405	small GTP-binding
20	506	45.8	216	2 S57471	GTP-binding protei
21	505.5	45.7	217	2 S36365	GTP-binding protei
22	502.5	45.5	215	2 S7462	GTP-binding protei
23	501	45.3	203	2 S14955	GTP-binding protei
24	495.5	44.8	204	2 J37589	Sec4p homolog - ye
25	492.5	44.6	215	2 S57474	GTP-binding protei
26	483.5	43.8	208	2 A34716	GTP-binding protei
27	482	43.7	208	2 A38202	GTP-binding protei
28	482	43.6	203	2 S34253	GTP-binding protei
29	480	43.4	202	2 S38740	GTP-binding protei

30	479	43.3	203	2 B34716	GTP-binding protei
31	478	43.3	202	2 J72515	GTP-binding protei
32	477.5	43.2	201	2 J2488	GTP-binding protei
33	477.5	43.2	201	2 J2488	GTP-binding protei
34	477.5	43.2	201	2 J2488	GTP-binding protei
35	477	43.2	202	2 J2488	GTP-binding protei
36	476	43.1	258	2 B86153	ARA-5 [imported] -
37	475.5	43.0	203	2 J2488	ARA-5 [imported] -
38	475	43.0	203	2 S30096	GTP-binding protei
39	473.5	42.9	205	2 T33781	hypothetical protei
40	473	42.8	206	2 T14391	GTP-binding protei
41	473	42.8	210	2 T18242	ras protein homolo
42	470.5	42.6	203	2 B38202	GTP-binding protei
43	470.5	42.6	215	1 TVBYQ4	GTP-binding protei
44	469	42.4	205	2 S38339	GTP-binding protei
45	468	42.4	205	1 TVRTYP	GTP-binding protei

ALIGNMENTS

RESULT 1

F42148

GTP-binding protein rab15 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 02-Feb-2001

C:Accession: F42148

R:Elferink, L.A.; Anzai, K.; Scheller, R.H.

J. Biol. Chem. 267, 5768-5775, 1992

A:Title: rab15, a novel low molecular weight GTP-binding protein specifically express

A:Reference number: A42148; MUID:92210533; PMID:1313420

A:Accession: F42148

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-212 <ELF>

C:Cross-references: GB:M33679; NID:g206536; PIDN:AAA41995.1; PID:g206537

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

C:Keywords: GTP binding; lipoprotein; nucleotide binding; P-loop; prenylated cysteine

F:9-124/Domain: translation elongation factor Tu homology <ETU>

F:15-22/Region: nucleotide-binding motif A (P-loop)

F:121-124/Region: GTP-binding NKXD motif

F:210,212/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 97.5%; Score 1077; DB 2; Length 212;

Best Local Similarity 97.6%; Pred. No. 9.2e-79;

Matches 207; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHSHSTIGVDFKMTIEVDGKVRIQ 60

Db 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHSHSTIGVDFKMTIEVDGKVRIQ 60

QY 61 IWDTAGQERYQITIKQYRRAGQIFLVYDIDISERSYQHIMKWSVDVAPGVOKILIG 120

Db 61 IWDTAGQERYQITIKQYRRAGQIFLVYDIDISERSYQHIMKWSVDVAPGVOKILIG 120

QY 121 NKADEQKRVGREGQQLAKEYGMDFYETSACTNLNIKESFTRTELVLQAHKLEGL 180

Db 121 NKADEQKRVGREGQQLAKEYGMDFYETSACTNLNIKESFTRTELVLQAHKLEGL 180

QY 181 RMRASNELALAELEEEKGPEGPANSSKTCWC 212

Db 181 RTCASNELALAELEDEGKTEGPANSSKTCWC 212

RESULT 2

T33855

hypothetical protein D1037.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 11-Jan-2000

C:Accession: T33855

R:Ledwith, J.; Biewald, T.

A:Description: The sequence of C. elegans cosmid D1037.

A:Reference number: Z21424

A:Accession: T33855

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-224 <LED>

A:Cross-references: EMBL:AF106592; PIDN:AAC78494.1; GSPDB:GN00019; CESP:D1037.4

A:Experimental source: strain Bristol N2; clone D1037

C:Genetics:

A:Gene: CESP:D1037.4

A:Map position: 1

A:Introns: 10/3; 62/2; 82/3; 181/1

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

Query Match

Best Local Similarity 50.6%; Score 559; DB 2; Length 224;

Matches 105; Conservative 42; Mismatches 39; Indels 4; Gaps 1;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSTIGVDFKMKTIEVDGKVRQ 60

Db 1 MAKTYDLFKLLIGDSGVGKTCVLFSEDFNSFISTIGIDFKIRTIELDGKRIKIQ 60

QY 61 IWDTAGQERYQTITKQYRRAQGIPLVYDIDISSERSYQHIMKMWSDVEYAPGVOKILIG 120

Db 61 IWDTAGQERFTITTAIRYGAMGIMLVYDITNEKSFENIKWIRNIEHSAADVEKMIIG 120

QY 121 NKADDEQKROVREGOGQOLAKYGMDFYETSACTNINIKESFTRTELVLQAHKRELEGL 180

Db 121 NKCDIEERREVRDRGEQLAIEYGTKEFLETSKANINIDENAFITLARDI---KSKMEQN 176

QY 181 RMRASNE 187

Db 177 EMRAATGA 186

RESULT 3

B49647

GTP-binding protein rab8 - human

C:Species: Homo sapiens (man)

C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 19-Jan-2001

C:Accession: B49647; S36817

R:Zahraoui, A.; Joberty, G.; Arpin, M.; Fontaine, J.J.; Hellio, R.; Tavittian, A.; Louvar

J. Cell Biol. 124, 101-115, 1994

A:Title: A small rab GTPase is distributed in cytoplasmic vesicles in non polarized cell

A:Reference number: A49647; MUID:94124602; PMID:8294494

A:Accession: B49647

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-207 <ZAH>

A:Cross-references: EMBL:X56741; NID:g452317; PIDN:CAA40065.1; PID:g452318

R:Joberty, G.; Tavittian, A.; Zahraoui, A.

FERS Lett. 330, 323-328, 1993

A:Title: Isoprenylation of Rab proteins possessing a C-terminal Caax motif.

A:Reference number: S36817; MUID:93387463; PMID:8375503

A:Accession: S36817

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 175-186 <JOB>

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

C:Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop;

F:9-124/Domain: translation elongation factor Tu homology <ETU>

F:15-22/Region: nucleotide-binding motif A (P-loop)

F:121-124/Region: translation elongation factor Tu homology

F:151-153/Region: GTP-binding NKXD motif

F:204/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match

Best Local Similarity 49.3%; Score 545; DB 2; Length 207;

Matches 99; Conservative 45; Mismatches 43; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSTIGVDFKMKTIEVDGKVRQ 60

Db 1 MAKTYDLFKLLIGDSGVGKTCVLFSEDFNSFISTIGIDFKIRTIELDGKRIKIQ 60

QY

61 IWDTAGQERYQTITKQYRRAQGIPLVYDIDISSERSYQHIMKMWSDVEYAPGVOKILIG 120

Db 61 IWDTAGQERFTITTAIRYGAMGIMLVYDITNEKSFENIKWIRNIEHSAADVEKMIIG 120

QY 121 NKADDEQKROVREGOGQOLAKYGMDFYETSACTNINIKESFTRTELVLQAHKRELEGL 180

Db 121 NKCDVNDKROVSKERGERKALDYGIKFMETSAKANINVENAFITLARDIKAKMDKKEGN 180

QY 181 RMRASNE 187

Db 181 SPQGSNQ 187

RESULT 4

B36364

GTP-binding protein rab8 - dog

C:Species: Canis lupus familiaris (dog)

C:Date: 28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change 02-Feb-2001

C:Accession: B36364; S15604

R:Chavrier, P.; Vingron, M.; Sander, C.; Simons, K.; Zerlial, M.

Mol. Cell Biol. 10, 6578-6585, 1990

A:Title: Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell line.

A:Reference number: A36364; MUID:91061765; PMID:2123294

A:Accession: B36364

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-207 <CHA>

A:Cross-references: GB:X56385; NID:g920; PIDN:CAB56776.1; PID:g6006436

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

C:Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop

F:9-124/Domain: translation elongation factor Tu homology

F:15-22/Region: nucleotide-binding motif A (P-loop)

F:121-124/Region: GTP-binding NKXD motif

F:151-153/Region: GTP-binding SAK/L motif

F:204/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match

Best Local Similarity 49.3%; Score 545; DB 2; Length 207;

Matches 99; Conservative 45; Mismatches 43; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSTIGVDFKMKTIEVDGKVRQ 60

Db 1 MAKTYDLFKLLIGDSGVGKTCVLFSEDFNSFISTIGIDFKIRTIELDGKRIKIQ 60

QY 61 IWDTAGQERYQTITKQYRRAQGIPLVYDIDISSERSYQHIMKMWSDVEYAPGVOKILIG 120

Db 61 IWDTAGQERFTITTAIRYGAMGIMLVYDITNEKSFENIKWIRNIEHSAADVEKMIIG 120

QY 121 NKADDEQKROVREGOGQOLAKYGMDFYETSACTNINIKESFTRTELVLQAHKRELEGL 180

Db 121 NKCDVNDKROVSKERGERKALDYGIKFMETSAKANINVENAFITLARDIKAKMDKKEGN 180

QY 181 RMRASNE 187

Db 181 SPQGSNQ 187

RESULT 5

A49647

GTP-binding protein Rab13 - human

C:Species: Homo sapiens (man)

C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 02-Feb-2001

C:Accession: A49647

R:Zahraoui, A.; Joberty, G.; Arpin, M.; Fontaine, J.J.; Hellio, R.; Tavittian, A.; Lou

J. Cell Biol. 124, 101-115, 1994

A:Title: A small rab GTPase is distributed in cytoplasmic vesicles in non polarized c

A:Reference number: A49647; MUID:94124602; PMID:8294494

A:Accession: A49647

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-203 <ZAH>

A:Cross-references: EMBL:X75593; NID:g452319; PIDN:CAA53266.1; PID:g452320

C:Superfamily: ras transforming protein; translation elongation factor Tu homology

Db	61	IWDTAGQERFRTITAYRGAMGIMKYVDITNEKSFNWKWIRNIESSHASSDVERMIIG	120
Qy	121	NKADEQKROVGREGOOLAKKEYGMDVFETSACTNLNIKESFTRLTETLVQAHRLEGL	180
Db	121	NKCDMNEKROVSKERGLADYGIK-ETSAKSSINVEEAFITLARDIMKLNKCM---	176

QY 161 RMRASNELALAB-LEEBEGKPEGPANSSK 208
 Db 177 -----NENSLQEAVDKSLPPKPSQKKK 200

RESULT 9

I78851
 GTP-binding protein MEL - mouse
 N:Alternate names: gene MEL protein
 C:Species: Mus sp. (mouse)
 C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 02-Feb-2001
 C:Accession: I78851
 C:Nummo, E.R.; Sanders, P.G.; Padua, R.A.; Hughes, D.; Williamson, R.; Johnson, K.J.
 Oncogene 6, 1347-1351, 1991
 A:Title: The MEL gene: a new member of the RAB/YTP class of RAS-related genes.
 A:Reference number: I58355; MUID:91360267; PMID:1886711
 A:Accession: I78851
 A:Status: Preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-206 <RES>
 A:Cross-references: GB:S53270; NID:9234747; PIDN:AAB19682.1; PID:9234748
 C:Genetics:
 A:Gene: MEL
 C:Superfamily: ras transforming protein; translation elongation factor Tu homology
 C:Keywords: GTP binding; lipoprotein; nucleotide binding; P-loop; prenylated cysteine
 F:9-124/Domain: translation elongation factor Tu homology <ETU>
 F:15-22/Region: nucleotide-binding motif A (P-loop)
 F:121-124/Region: GTP-binding NRXD motif
 F:151-153/Region: GTP-binding SAK/L motif
 F:203/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 47.6%; Score 526; DB 2; Length 206;
 Best Local Similarity 56.4%; Pred. No. 7.9e-35;
 Matches 93; Conservative 41; Mismatches 31; Indels 0; Gaps 0;
 QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSHISTIGVDFKMTIEVDGKVRQI 60
 Db 1 MAKTYDYLKLLIGDSGVGKTCVLFPSDADFNTTIGDFKIRTIELDKRRIKLQ 60
 QY 61 IWDTAGQERYQTITKQYRRAGQIFLVYDISERSYOHIMKWVSDVDEYAPGVOKILIG 120
 Db 61 IWDTAGQERFRTITAYYRGAMGIMLVYDITNEKSFDIRNWRNIEEHASADVKEMLIG 120
 QY 121 NKADEQKRVQREGQQLAKEYGMDFYETSACTNINIKESFTRL 165
 Db 121 NKCDVNDKRVQSKERGEKLADYGIKFMETSAKANINVENAFPTL 165

RESULT 10

B42148
 GTP-binding protein rab10 - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 02-Feb-2001
 C:Accession: B42148
 R:Elferink, L.A.; Anzal, K.; Scheller, R.H.
 J. Biol. Chem. 267, 5768-5775, 1992
 A:Title: rab15, a novel low molecular weight GTP-binding protein specifically expressed
 A:Reference number: A42148; MUID:92210533; PMID:1313420
 A:Accession: B42148
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-200 <ELF>
 A:Cross-references: GB:M83677
 C:Superfamily: ras transforming protein; translation elongation factor Tu homology
 C:Keywords: GTP binding; lipoprotein; nucleotide binding; P-loop; prenylated cysteine
 F:10-125/Domain: translation elongation factor Tu homology <ETU>
 F:16-23/Region: nucleotide-binding motif A (P-loop)
 F:122-125/Region: GTP-binding NRXD motif
 F:152-154/Region: GTP-binding SAK/L motif
 F:199,200/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 47.3%; Score 522.5; DB 2; Length 200;

Best Local Similarity 55.2%; Pred. No. 1.5e-34;
 Matches 95; Conservative 46; Mismatches 30; Indels 1; Gaps 1;
 QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSHISTIGVDFKMTIEVDGKVRQI 59
 Db 1 MAKTYDYLKLLIGDSGVGKTCVLFPSDADFNTTIGDFKIRTIELDKRRIKLQ 60
 QY 60 QIWDTAGQERYQTITKQYRRAGQIFLVYDISERSYOHIMKWVSDVDEYAPGVOKILI 119
 Db 61 QIWDTAGQERFRTITAYYRGAMGIMLVYDITNGSFENISKWLNIHANEDVERML 120
 QY 120 GNKADEQKRVQREGQQLAKEYGMDFYETSACTNINIKESFTRLTELVLQ 171
 Db 121 GNKCDMDKRVVPKGGQIAREHGIRFPFETSAKANINIEKAFLLAEDILR 172

RESULT 11

T45901
 GTPase ATRAB8 - Arabidopsis thaliana
 N:Alternate names: protein F4P12.310
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 02-Sep-2000
 C:Accession: T45901
 R:Blöcker, H.; Mewes, H.W.; Lencke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat, M.
 submitted to the Protein Sequence Database, January 2000
 A:Reference number: Z23016
 A:Accession: T45901
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-216 <BLQ>
 A:Cross-references: EMBL:AL132966
 A:Experimental source: cultivar Columbia; BAC clone F4P12
 C:Genetics:
 A:Map position: 3
 A:Introns: 25/1; 49/2; 83/1; 105/1; 137/3; 158/3; 189/3
 A:Note: F4P12.310
 C:Superfamily: ras transforming protein; translation elongation factor Tu homology
 F:16-131/Domain: translation elongation factor Tu homology <ETU>

Query Match 47.0%; Score 519; DB 2; Length 216;

Best Local Similarity 51.9%; Pred. No. 3e-34;
 Matches 98; Conservative 42; Mismatches 43; Indels 6; Gaps 2;
 QY 5 YDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSHISTIGVDFKMTIEVDGKVRQIWD 64
 Db 12 YDYLKLLIGDSGVGKSCLLRFRSDGSFTTIFITIGDFKIRTIELDKRRIKLQIWD 71
 QY 65 AGQERYQTITKQYRRAGQIFLVYDISERSYOHIMKWVSDVDEYAPGVOKILIGNKAD 124
 Db 72 AGQERFRTITAYYRGAMGIMLVYDITDESFNNRNWRNIEEHASGVNKLIGNKAD 131
 QY 125 -EEQKRVQREGQQLAKEYGMDFYETSACTNINIKESFTRLTELVLQA-----HRKELE 178
 Db 132 MDESKRAVPKSGQALADEYGMKPFETSAKTNINVEEVFFSTAKDIKQRLADYDARAEPQ 191
 QY 179 GLRMPASNE 187
 Db 192 TIKINQSDQ 200

RESULT 12

T28971
 hypothetical protein T23H2.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-Jan-2000
 C:Accession: T28971
 R:Wansley, P.; Bradshaw, H.
 submitted to the EMBL Data Library, January 1997
 A:Description: The sequence of C. elegans cosmid T23H2.
 A:Reference number: Z20549
 A:Accession: T28971
 A:Status: Preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA


```

A;Residues: 1-201 <WAM>
A;Cross-references: EMBL:U80033; PIDN:AAC48200.1; GSPDB:GN00019; CESP:T23H2.5
A;Experimental source: strain Bristol N2; clone T23H2
C;Genetics:
  A;Gene: CESP:T23H2.5
  A;Map position: 1
  A;Introns: 43/1; 147/2; 173/3
  C;Superfamily: ras transforming protein; translation elongation factor Tu homology

Query Match          46.5%; Score 514; DB 2; Length 201;
Best Local Similarity 46.9%; Pred. No. 7e-34;
Matches 100; Conservative 48; Mismatches 51; Indels 14; Gaps 3;

QY   1 MAKQ-YDVLFRLLIGDSGVGKTCCLCRDTNEFHSHISTIGVDFKMKTIEVDGIKVRI 59
      ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::|
Db    1 MARRPYDMFLKLLIGDSGVGTCLILRPSDDAFNTFTSTICIDFKIKTIELKGKKIKL 60

QY   60 QIWDTAGQRYSOTITKQYYRAAGIFLVYDISSRSYQHIMKWSDVDVEYAPGVOKILI 119
      ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::|
Db    61 QIWDTAGQRFFHTITSYIRGAWGLIMLYDIINAKSFDNIKLWRINDSHASDVVKMIL 120

QY   120 GNKADEQRQVRGEQGQOLAKEYGMDFYETSACTNLNIKESFTRLTVELVLAHRKELEG 179
      ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::|
Db    121 GNKCMSDRRVSRERGERIAQDHGISFTHETSAKLNHVHYDTAFYLAEAIL----- 171

QY   180 LMRASNELALALEEBEKEGPGPANSSKTCWC 212
      :||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db    172 AKMPDSTD---EQSRDTVPNPVQPQRQSSGGC 200

RESULT 13
S33900
GTP-binding protein ypt2 - tomato
C;Species: Lycopersicon esculentum (tomato)
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 02-Feb-2001
C;Accession: S33900; J02233
R;Fleming, A.J.; Mandel, T.; Roth, I.; Kuhlemeier, C.
Plant Cell 5, 297-309, 1993
A;Title: The patterns of gene expression in the tomato shoot apical meristem.
A;Reference number: S33899; MUID:93222691; PMID:8467223
A;Accession: S33900
A;Molecule type: mRNA
A;Residues: 1-216 <PL2>
A;Cross-references: GB:X69980; NID:g313028; PIDN:CAA49600.1; PID:g313029
C;Genetics:
  A;Gene: YPT2
  C;Superfamily: ras transforming protein; translation elongation factor Tu homology
  C;Keywords: GTP binding; membrane protein; nucleotide binding; P-loop
  F;16-131/Domain: translation elongation factor Tu homology <ETU>
  F;22-29/Region: nucleotide-binding motif A (P-loop)
  F;128-131/Region: GTP-binding NKAD motif
  F;159-161/Region: GTP-binding SAK/L motif

Query Match          46.3%; Score 512; DB 2; Length 216;
Best Local Similarity 50.5%; Pred. No. 1.1e-33;
Matches 99; Conservative 41; Mismatches 44; Indels 12; Gaps 3;

QY   5 YDVLFRLLIGDSGVGKTCCLCRDTNEFHSHISTIGVDFKMKTIEVDGIKVRIQIWDT 64
      ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::|
Db    12 YDYLKLLLLIGTGVGSKLLLRFSDFSFTTFITTIGIDFKIRTIELDGKRKLKIWDWT 71

QY   65 AGQERYQTITKQYYRAAGIFLVYDISSRSYQHIMKWSDVDVEYAPGVOKILIGNKAD 124
      ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::|
Db    72 AGQERFTTTAYRGAGILIVDYDTSSEFNNTKNRNIEHQHASDNVNKILVGNKAD 131

QY   125 -EQKRVQREGQQQLAKEYGMDFYETSACTNLNIKESF-----TRLTELVLQAHRK 175
      :||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db    132 MDKSRAVPTSQGALEADYEGIKFFTSAKTLNVNEEVFFSGDKDKIQRLSE---SDSKT 188

QY   176 ELEGLRMNASNELALA 191
Db    189 EPQSIRINQSDQAGTA 204

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:24:36 ; Search time 22 Seconds
(without alignments)
399.681 Million cell updates/sec

Title: US-09-817-198B-2

Perfect score: 1105

Sequence: 1 MAKQYDVLFRLLIGSGVG.....LEEEKGPEGPANSSKTCWC 212

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1077	97.5	212	1 RB15_RAT	P35289 rattus norv
2	545	49.3	207	1 RAB8_HUMAN	P24407 homo sapien
3	539.5	48.8	210	1 RAB8_DISOM	P22128 discopoyge o
4	532	48.1	203	1 RB13_HUMAN	P51153 homo sapien
5	531.5	48.1	200	1 RAOL_DISOM	P22127 discopoyge o
6	530.5	48.0	200	1 RB10_CANFA	P24409 canis famil
7	530.5	48.0	200	1 RB10_HUMAN	O88386 homo sapien
8	528.5	47.8	207	1 RB8B_HUMAN	Q92930 homo sapien
9	528.5	47.8	207	1 RB8B_RAT	P70530 rattus norv
10	526	47.6	206	1 RAB8_MOUSE	P55258 mus musculu
11	510	46.2	200	1 YPT2_SCHPO	P17609 schizosacch
12	509	46.1	215	1 RAB1_BETVU	Q39433 beta vulgar
13	508	46.0	216	1 ARA3_ARATH	P28186 arabidopsis
14	505.5	45.7	217	1 YPT2_VOLCA	P36861 volvox cart
15	502.5	45.5	200	1 RB10_RAT	P35281 rattus norv
16	502	45.4	203	1 RYLL_YARLI	P41924 yarrowia li
17	483.5	43.8	208	1 SAS1_DICDI	P20790 dictyosteli
18	483	43.7	208	1 YPT1_MAZE	P16976 zea mays (m
19	480	43.4	202	1 RIC1_ORYSA	P40392 oryza sativ
20	479	43.3	203	1 SAS2_DICDI	P20791 dictyosteli
21	477.5	43.2	201	1 RB35_HUMAN	Q15286 homo sapien
22	477.5	43.2	203	1 YPT1_CHLRE	Q39571 chlamydomon
23	477	43.2	201	1 YPT1_PHYIN	Q01890 phytophthor
24	476	43.1	258	1 ARA5_ARATH	P28188 arabidopsis
25	475.5	43.0	203	1 YPT1_VOLCA	P31584 volvox cart
26	475	43.0	203	1 YPT1_NEUCR	P33723 neurospora
27	473	42.8	210	1 SEC4_CANAL	Q14462 candida alb
28	470.5	42.6	203	1 YPT2_MAZE	Q05737 zea mays (m
29	470.5	42.6	215	1 SEC4_YEAST	P07560 saccharomyc
30	469	42.4	201	1 RB1B_RAT	P10536 rattus norv
31	469	42.4	205	1 RAB1_LYMST	Q05974 lymnaea sta
32	468	42.4	205	1 RB1A_RAT	P05711 rattus norv
33	467	42.3	203	1 YPT1_SCHPO	P11620 schizosacch

ALIGNMENTS

RESULT_1

ID	RB15_RAT	STANDARD;	PRT;	212 AA.
AC	P35289;			
DT	01-FEB-1994 (Rel. 28, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	01-OCT-1996 (Rel. 34, Last annotation update)			
DE	Ras-related protein Rab-15.			
GN	RAB15.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A. .			
RC	STRAIN-Sprague-Dawley; TISSUE-Brain;			
RX	MEDLINE=92210533; PubMed=1313420;			
RA	Elferink L.A., Anzai K., Scheller R.H.;			
RT	"Rab15, a novel low molecular weight GTP-binding protein specifically expressed in rat brain.";			
RL	J. Biol. Chem. 267:5768-5775(1992).			
RN	[2]			
RP	ERRATUM.			
RX	MEDLINE=93054572; PubMed=1429617;			
RA	Elferink L.A., Anzai K., Scheller R.H.;			
RL	J. Biol. Chem. 267:22693-22693(1992).			
CC	-1- FUNCTION: MAY ACT IN CONCERT WITH RAB3A IN REGULATING ASPECTS OF SYNAPTIC VESICLE MEMBRANE FLOW WITHIN THE NERVE TERMINAL.			
CC	-1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN NEURAL TISSUES.			
CC	-1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; M83679; AAA41995.1; -			
DR	PIR; F42148; F42148.			
DR	HSSP; P05713; 3RAB.			
DR	InterPro; IPR003579; GTPase_Rab.			
DR	InterPro; IPR001806; Ras_trnsfrmng.			
DR	InterPro; IPR005225; Small_GTP.			
DR	Pfam; PF00071; ras; 1.			
DR	PRINTS; PR00449; RASTRNSFRMNG.			
DR	SMART; SM00175; RAB; 1.			
DR	TIGRFAMS; TIGR00231; small_GTP; 1.			
KW	GTP-binding; Lipoprotein; Prenylation; Protein transport.			
FT	NP_BIND 15 22 GTP (BY SIMILARITY).			
FT	NP_BIND 63 67 GTP (BY SIMILARITY).			
FT	NP_BIND 121 124 GTP (BY SIMILARITY).			
FT	LIPID 210 210 GERANYL-GERANYL (BY SIMILARITY).			
FT	LIPID 212 212 GERANYL-GERANYL (BY SIMILARITY).			
SQ	SEQUENCE 212 AA; 24283 MW; 04817DDA66CADE12 CRC64;			

P11476 homo sapien
P22125 discopoyge o
P01123 saccharomyc
Q63482 rattus norv
Q9ulw5 homo sapien
P51156 rattus norv
P25228 drosophila
O95716 homo sapien
P11023 bos taurus
O63942 rattus norv
P05713 mus musculu
P20336 homo sapien

```

Query Match      97.5%; Score 1077; DB 1; Length 212;
Best Local Similarity 97.6%; Pred. No. 6.9e-80;
Matches 207; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSHSTIGVDPKMTIEVDGKVRQ 60
DB 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSHSTIGVDPKMTIEVDGKVRQ 60

QY 61 IWDTAGQRYQTITKQYRRAGGIFLVYDIDISSERSYQHIMKWSVDVEYAPGVOKILIG 120
DB 61 IWDTAGQRYQTITKQYRRAGGIFLVYDIDISSERSYQHIMKWSVDVEYAPGVOKILIG 120

QY 121 NKADEQRQVREGOQOLAKKEYGMDFYETSACTNINIKESFTRTLVLQAHKRLGL 180
DB 121 NKADEQRQVREGOQOLAKKEYGMDFYETSACTNINIKESFTRTLVLQAHKRLGL 180

QY 181 RMRASNELALAELEEGKPEGPANSSKTCWC 212
DB 181 RMRASNELALAELEEGKPEGPANSSKTCWC 212

QY 181 RMRASNELALAELEEGKPEGPANSSKTCWC 212
DB 181 RMRASNELALAELEEGKPEGPANSSKTCWC 212

RESULT 2
RAB8_HUMAN
ID RAB8_HUMAN STANDARD; PRT; 207 AA.
AC P24407;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Ras-related protein Rab-8 (Oncogene c-mel).
GN RAB8 OR MEL.
OS Homo sapiens (Human), and
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606, 9615;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Human;
RX MEDLINE=94124602; PubMed=8994494;
RA Zahraoui A., Joberty G., Arpin M., Fontaine J.J., Hellio R.,
RA Tavittian A., Louvard D.;
RT "A small rab GTPase is distributed in cytoplasmic vesicles in non
RT polarized cells but colocalizes with the tight junction marker ZO-1
RT in polarized epithelial cells.";
RL J. Cell Biol. 124:101-115(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Human;
RX MEDLINE=91360267; PubMed=1886711;
RA Nimmo E.R., Sanders P.G., Padua R.A., Hughes D., Williamson R.,
RA Johnson K.J.;
RT "The MEL gene: a new member of the RAB/YPT class of RAS-related
RT genes.";
RL Oncogene 6:1347-1351(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=Human;
RX Strausberg R.;
RA Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=C.familiaris; STRAIN=Cocker spaniel;
RX MEDLINE=91061765; PubMed=2123294;
RA Chavrier P., Vingron M., Sander C., Simons K., Zerial M.;
RT "Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell
RT line.";
RL Mol. Cell. Biol. 10:6578-6585(1990).
CC -!- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND
CC NEUROTRANSMITTER RELEASE.
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X56741; CAA40065.1; -
DR EMBL; S53268; AAB19681.1; -
DR EMBL; BC002977; AA02977.1; -
DR EMBL; X56385; CAB56776.1; -
DR PIR; B36364; B36364.
DR PIR; B49647; B49647.
DR HSSP; P05713; 3RAB.
DR GENE; HGNC:7007; MEL.
DR MIM; 165040; -
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001230; Prenyl_site.
DR InterPro; IPR001806; Ras_transfmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR TIGRFS; TIGR00231; small_GTP; 1.
KW GTP-binding; Prenylation; Lipoprotein; Protein transport;
KW Proto-oncogene.
FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 121 124 GTP (BY SIMILARITY).
FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 204 204 GERANYL-GERANYL (BY SIMILARITY).
FT CONFLICT 177 183 LEGNSPQ -> WKATAP (IN REF. 2).
SQ SEQUENCE 207 AA; 23668 MW; AA52DBF5A2CD056 CRC64;

Query Match      49.3%; Score 545; DB 1; Length 207;
Best Local Similarity 52.9%; Pred. No. 4.2e-37;
Matches 99; Conservative 45; Mismatches 43; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSHSTIGVDPKMTIEVDGKVRQ 60
DB 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSHSTIGVDPKMTIEVDGKVRQ 60

QY 61 IWDTAGQRYQTITKQYRRAGGIFLVYDIDISSERSYQHIMKWSVDVEYAPGVOKILIG 120
DB 61 IWDTAGQRYQTITKQYRRAGGIFLVYDIDISSERSYQHIMKWSVDVEYAPGVOKILIG 120

QY 121 NKADEQRQVREGOQOLAKKEYGMDFYETSACTNINIKESFTRTLVLQAHKRLGL 180
DB 121 NKADEQRQVREGOQOLAKKEYGMDFYETSACTNINIKESFTRTLVLQAHKRLGL 180

QY 181 RMRASNE 187
DB 181 SPQGSNQ 187

RESULT 3
RAB8_DISOM
ID RAB8_DISOM STANDARD; PRT; 210 AA.
AC P22128;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Ras-related protein Rab-8 (ORA2).
OS Discopoge omata (Electric ray).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Hypnosqualea; Pristiorajae; Batoidae;
OC Torpediniformes; Narcinoidei; Narcinidae; Discopogae.
OX NCBI_TaxID=7785;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Electric lobe;
RX MEDLINE=91115900; PubMed=1899244;
RA Ngsee J.K., Elferink L.A., Scheller R.H.;

```

"A family of ras-like GTP-binding proteins expressed in electromotor neurons.";
 J. Biol. Chem. 266:2675-2680(1991).
 -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

 EMBL; M38391; AAA49232.1; --
 PIR; B38625; B38625.
 HSP; P05713; 3RAB.
 InterPro; IPR003579; GTPase_Rab.
 InterPro; IPR001230; Prenyl_site.
 InterPro; IPR001806; Ras_trnsmfmg.
 InterPro; IPR005225; Small_GTP.
 Pfam; PF00071; ras; 1.
 PRINTS; PR00449; RASTRNSFRMNG.
 SMART; SM00175; RAB; 1.
 TIGREMS; TIGR00231; small_gtp; 1.
 GTP-binding; Prenylation; Lipoprotein.
 NP_BIND 15 22 GTP (BY SIMILARITY).
 NP_BIND 63 67 GTP (BY SIMILARITY).
 NP_BIND 121 124 GTP (BY SIMILARITY).
 DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).
 LIPID 207 207 GERANYL-GERANYL (BY SIMILARITY).
 SEQUENCE 210 AA; 24164 MW; 0F73EDB0DB9BEA CRC64;

 Query Match 48.8%; Score 539.5; DB 1; Length 210;
 Best Local Similarity 47.4%; Pred. No. 1.2e-36;
 Matches 99; Conservative 56; Mismatches 45; Indels 9; Gaps 2;

 QY 1 MAKQYDLVFLRLIGDSGVGKTCCLCRFTDNEFHSHISTIGVDFKMTIEVDGKVRQ 60
 DB 1 MAKYDYLFKLLIGDSGVGKTCCLIRFAEDFNNTYISTIGIDFKIRTVDEGKKIKLQ 60

 QY 61 IWDTAGQERYQTITKQYRRAQGIPLVYDIDISSERSYQHIMKWSVDVEYAPGQKILIG 120
 DB 61 IWDTAGQERFTITAYRGAMGIMKYVDITNEKSFNIRWIRNIEHASSDVERMILG 120

 QY 121 NKADEQKRVGRQGOOLAKEYGMDFYETSACTNLIKESFTLRLTLVLQAHKRLGL 180
 DB 121 NKCDMEKRVQSKERGEKLAIDYGIKEFLETSKSSINVEAFITLARDIMTKLNKM-- 177

 QY 181 RMRASNELALAE-LEEEEGKPEGPANSSK 208
 DB 178 -----NENSLQEAVDKILKSPKPSQKK 201

 RESULT 4
 RB13_HUMAN
 ID RB13_HUMAN STANDARD; PRT; 203 AA.
 AC P51153;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ras-related protein Rab-13.
 GN RB13.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94124602; PubMed=8294494;
 RA Zahraoui A., Joberty G., Arpin M., Fontaine J.J., Hellio R.,
 RA Tavittian A., Louvard D.;
 RT "A small rab GTPase is distributed in cytoplasmic vesicles in non
 polarized cells but colocalizes with the tight junction marker ZO-1

in polarized epithelial cells.";
 J. Cell Biol. 124:101-115(1994).
 [2]
 SEQUENCE FROM N.A.
 TISSUE-Placenta;
 Submitted R.;
 Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 -1- FUNCTION: COULD PARTICIPATE IN POLARIZED TRANSPORT, IN THE ASSEMBLY AND/OR THE ACTIVITY OF TIGHT JUNCTIONS.
 -1- SUBCELLULAR LOCATION: CYTOPLASMIC TIGHT JUNCTIONS OR ASSOCIATED WITH VESICLES SCATTERED THROUGHOUT THE CYTOPLASM IN CELLS LACKING TIGHT JUNCTIONS.
 -1- TISSUE SPECIFICITY: VARIETY OF EPITHELIA, INCLUDING INTESTINE, KIDNEY, LIVER, AND OF ENDOTHELIAL CELLS.
 -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

 EMBL; X75593; CAA53266.1; --
 EMBL; BC000799; AAH00799.1; --
 HSP; P01112; 1PLL.
 Genew; HGNC:9762; RAB13.
 MIM; 602672; --
 InterPro; IPR003579; GTPase_Rab.
 InterPro; IPR001230; Prenyl_site.
 InterPro; IPR001806; Ras_trnsmfmg.
 InterPro; IPR005225; Small_GTP.
 Pfam; PF00071; ras; 1.
 PRINTS; PR00449; RASTRNSFRMNG.
 SMART; SM00175; RAB; 1.
 TIGREMS; TIGR00231; small_gtp; 1.
 GTP-binding; Lipoprotein; Prenylation; Protein transport.
 NP_BIND 15 22 GTP (BY SIMILARITY).
 NP_BIND 63 67 GTP (BY SIMILARITY).
 NP_BIND 121 124 GTP (BY SIMILARITY).
 DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).
 LIPID 200 207 GERANYL-GERANYL (BY SIMILARITY).
 SEQUENCE 203 AA; 22774 MW; 14162LCB998178DA CRC64;

 Query Match 48.1%; Score 532; DB 1; Length 203;
 Best Local Similarity 47.6%; Pred. No. 4.6e-36;
 Matches 100; Conservative 49; Mismatches 43; Indels 18; Gaps 3;

 QY 1 MAKQYDLVFLRLIGDSGVGKTCCLCRFTDNEFHSHISTIGVDFKMTIEVDGKVRQ 60
 DB 1 MAKYDYLFKLLIGDSGVGKTCCLIRFAEDFNNTYISTIGIDFKIRTVDEGKKIKLQ 60

 QY 61 IWDTAGQERYQTITKQYRRAQGIPLVYDIDISSERSYQHIMKWSVDVEYAPGQKILIG 120
 DB 61 IWDTAGQERFTITAYRGAMGILYDIDDEKSFENIQNMKSIRKENSAGVERLLLG 120

 QY 121 NKADEQKRVGRQGOOLAKEYGMDFYETSACTNLIKESFTLRLTLVLQAHKRLGL 180
 DB 121 NKCDMEKRVQKQKQADKLAHEGIRFEFETSAKSMNVDEAFSSLDAIL-----LKSQ 174

 QY 181 RMRASNELALAELEEEEGKPEGPANSSKTC 210
 DB 175 GRRSGN-----GNKP--PSTDLKTC 192

 RESULT 5
 RA01_DISOM
 ID RA01_DISOM STANDARD; PRT; 200 AA.
 AC P22127;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 01-AUG-1991 (Rel. 19, Last annotation update)

RA Gu B.-W., Dai M., Mao Y.-F., Gao G.-F., Rong R., Ye M., Zhou J.,
 RA Xu S.-H., Gu J., Shi J.-X., Jin W.-R., Zhang C.-K., Wu T.-M.,
 RA Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.,
 RT *Gene expression profiling in the human hypothalamus-pituitary-adrenal
 RT axis and full-length cDNA cloning.*;
 RN Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).
 [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Human;
 RA Wong K., Hong W., Tang B.;
 RA Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Human;
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Watanabe M., Hosokawa T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yanamoto J., Wakamatsu A., Nakamura Y., Negahari K., Masuho Y.,
 RA Niinomiya K., Iwayanagi T.;
 RT *NDO human cDNA sequencing project.*;
 RT Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Human; TISSUE=Cervix;
 RA Strausberg R.;
 RA Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 [5]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Mouse;
 RA Zeng Q., Tan Y.H., Hong W.;
 RA Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 [6]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Mouse; STRAIN=C57BL/6J; TISSUE=Stomach;
 RX MEDLINE-21085660; PubMed-11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Akawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirni L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT *Functional annotation of a full-length mouse cDNA collection.*;
 RL Nature 409:685-690(2001).
 CC -1- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND
 CC NEUROTRANSMITTER RELEASE.
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF106681; AAD43034.1; -
 CC EMBL; AF297660; AAG13413.1; -
 CC EMBL; AK023223; BAB14474.1; -
 CC EMBL; BC000896; AAB00896.1; -

DR EMBL; AF035646; AAC29313.1; -
 DR EMBL; AK008725; BAB25858.1; -
 DR HSSP; P05713; 3RAB.
 DR Genew; HGNC:9759; RAB10.
 DR MGD; MGI:105066; RAB10.
 DR InterPro; IPR003579; GTPase_Rab.
 DR InterPro; IPR001806; Ras_trnsfrrng.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00071; ras; 1.
 DR PRINTS; PR00449; RASTRNSFRMG.
 DR SMART; SM00175; RAB; 1.
 DR TIGRFAMS; TIGR00231; small_GTP; 1.
 KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
 FT NP_BIND 16 23 GTP (BY SIMILARITY).
 FT NP_BIND 64 68 GTP (BY SIMILARITY).
 FT NP_BIND 122 125 GTP (BY SIMILARITY).
 FT DOMAIN 38 46 EFFECTOR REGION (BY SIMILARITY).
 FT LIPID 199 199 GERANYL-GERANYL (BY SIMILARITY).
 FT LIPID 200 200 GERANYL-GERANYL (BY SIMILARITY).
 FT CONFLICT 106 106 N -> H (IN REF. 6).
 SQ SEQUENCE 200 AA; 22541 MW; 7F02B8E8E46EE1E8 CRC64;
 Query Match 48.0%; Score 530.5; DB 1; Length 200;
 Best Local Similarity 55.8%; Pred. No. 5.9e-36;
 Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;
 QY 1 MAKQ-YDVLFRLLIGDSGVGKTCCLCRFTDNEPHSHSHISTIGYDFRMKTIETVDGIKVI 59
 DB 1 MAKTYDLLFKLLIGDSGVGKTCVLPFRFSDAFTFTFISTIGIDFKIKIVELQKKIKL 60
 QY 60 QIWDTAGERYQITIKOYIRRAQGIIFLYVIDISSRQIHKMKWSDVDEVAPEGVQKILI 119
 DB 61 QIWDTAGOERFHTITTYIRGAMGIMLVYDITNGKSPENISKWLRNIDEHANEDVERMLL 120
 QY 120 GNRADSEQKRVGREGQOLAKYGMDFYETASCTNLNKEISFRLRLVLIQ 171
 DB 121 GNKCDMDKRVVPKGEQIAREHGIRFFETSAKANINIEKAFLLAEDILR 172
 RESULT 8
 RB8B_HUMAN STANDARD; PRT; 207 AA.
 ID Q92930; Q9P293;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ras-related protein Rab-8B.
 GN RAB8B
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RA Seki N., Saito T.;
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RA Strausberg R.;
 RN Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 [3]
 RN SEQUENCE OF 67-119 FROM N.A.
 RP TISSUE=Melanoma;
 RX MEDLINE-97182150; PubMed-9030196;
 RA Chen D., Guo J., Gahl W.A.;
 RT "RAB GTPases expressed in human melanoma cells.";
 RL Biochim. Biophys. Acta 1355:1-6(1997).
 CC -1- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND
 CC NEUROTRANSMITTER RELEASE (BY SIMILARITY).
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

CC	ENBL; AB038995; BAA92249.1; -	
DR	ENBL; BC020654; AAH20654.1; -	
DR	ENBL; U66624; AAC51199.1; -	
DR	HSP; P05713; 3RAB.	
DR	InterPro; IPR003579; GTPase_Rab.	
DR	InterPro; IPR001230; Prenyl_site.	
DR	InterPro; IPR001806; Ras_trnsfrmng.	
DR	InterPro; IPR005225; Small_GTP.	
DR	Pfam; PF00071; ras; 1.	
DR	PRINTS; PR00449; RASTRNSFRMNG.	
DR	SMART; SM00175; RAB; 1.	
DR	TIGRams; TIGR00231; small_GTP; 1.	
KW	GTP-binding; Prenylation; Lipoprotein; Protein transport.	
FT	NP_BIND 15 22	GTP (BY SIMILARITY).
FT	NP_BIND 63 67	GTP (BY SIMILARITY).
FT	NP_BIND 121 124	GTP (BY SIMILARITY).
FT	DOMAIN 37 45	EFFECTOR REGION (BY SIMILARITY).
FT	LIPID 204 204	GERANYL-GERANYL (BY SIMILARITY).
FT	SEQUENCE 207 AA: 23584 MW: 5960993CQF87F944 CRC64:	

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

DR EMBL: US3475; AAA99782.1; -

DR	HSSP:	P05713;	3RAB.
DR	InterPro:	IPR003579;	GTPase_Rab.
DR	InterPro:	IPR001230;	Prenyl_site.
DR	InterPro:	IPR001806;	Ras_Crnsfrmg.
DR	InterPro:	IPR005225;	Small_GTP.
DR	Pfam:	PF00071;	ras; 1.
DR	PRINTS:	PR00449;	RASTRNSFRMG.
DR	SMART:	SM00175;	RAB; 1.
DR	TIGRAMS:	TIGR00231;	small_GTP; 1.
KW	GTP-binding;	Prenylation;	Lipoprotein; Protein transport.
FT	NP_BIND	15	22 GTP (BY SIMILARITY).
FT	NP_BIND	63	67 GTP (BY SIMILARITY).
FT	NP_BIND	121	124 GTP (BY SIMILARITY).
FT	DOMAIN	37	45 EFFECTOR REGION (BY SIMILARITY).
FT	LIPID	204	GERANYL-GERANYL (BY SIMILARITY).
SQ	SEQUENCE	207 AA;	23603 MW; A4A1AB26BF9DCAFA4 CRC64;
	Query Match		47.8%; Score 528.5; DB 1; Length 207;
	Best Local Similarity	54.0%;	Pred. No. 8.9e-36;
Matches	95: Conservative	47:	Mismatches 33: Indels 1: Gaps 1:

```

CC EMBL: S53270; AAB19682.1;
DR HSP: P05713; 3RAB.
DR MGD: MGI:96960; Mel.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001230; Prenyl_site.
DR InterPro: IPR001806; Ras_trnsfrmng.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; ras; 1.
DR PRINTS: PR00449; RASTRNSFRMG.
DR SMART: SM00175; RAB; 1.
DR TIGRFAMs: TIGR00231; small_GTP; 1.
KW GTP-binding; Prenylation; Lipoprotein; Protein transport;
KW Proto-oncogene.
FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 121 124 GTP (BY SIMILARITY).
FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 203 203 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 206 AA; 23557 MW; 49D832725D662942 CRC64;

Query Match 47.6%; Score 526; DB 1; Length 206;
Best Local Similarity 56.4%; Pred. No. 1.4e-35;
Matches 93; Conservative 41; Mismatches 31; Indels 0; Gaps 0;

QY 1 MAKQYDLVRLLLIGDSGVGKTCILCRFTDNEFHSHSTIGVDPKMTIEVDGIKVRIO 60
DB 1 MAKTYDLVFKLLIGDSGVGKTCVLFSEDAFNSTFTIGIDPKIRTIELDGKRIKIQ 60
QY 61 IWDTAGQERYQITIKQYRRAGQIFLVYDISERSYQIHMKVSDVDVAPGVGKILIG 120
DB 61 IWDTAGQERYFRITITAYYRGAMGIMLYDITNEKSFDRINRWIRNEHASADVEKMILG 120
QY 121 NKADEQKRVQREGQQLAKEYGMDFYETSACTNLNKESFTRL 165
DB 121 NKCDVNDKRVQSKERGEKIALDYGKFMETSAKANINVENAPFTL 165

RESULT 11
YPT2_SCHPO STANDARD; PRT: 200 AA.
AC P17609;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ras-related protein ypt2.
GN YPT2 OR SPAC9E9.07C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=90269232; PubMed=2112089;
RA Hengst L., Lemmer T., Gallwitz D.;
RT "Structural and functional analysis of ypt2, an essential ras-related
RT gene in the fission yeast Schizosaccharomyces pombe encoding a Sec4
RT protein homologue."
RL EMBO J. 9:1957-1962(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90332438; PubMed=2115995;
RA Fawell E., Hook S., Sweet D., Armstrong J.;
RT "Novel ypt1-related genes from Schizosaccharomyces pombe."
RL Nucleic Acids Res. 18:4264-4264(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,

```

```

RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Wambutt R., Purnelle B.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Sherratt L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -!- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
CC -!- TRAFFIC (BY SIMILARITY).
CC -!- MISCELLANEOUS: THIS PROTEIN IS ESSENTIAL FOR CELL VIABILITY.
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC PROBABLE YEAST SEC4 HOMOLOG.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X52469; CAA36707.1; -
DR EMBL: X52864; CAA37045.1; -
DR EMBL: Z99262; CABI6405.1; -
DR PIR: S10493; S10493.
DR PIR: S12790; S12790.
DR HSP: P05713; 3RAB.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_trnsfrmng.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; ras; 1.
DR PRINTS: PR00449; RASTRNSFRMG.
DR SMART: SM00175; RAB; 1.
DR TIGRFAMs: TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
FT NP_BIND 16 23 GTP (BY SIMILARITY).
FT NP_BIND 64 68 GTP (BY SIMILARITY).
FT NP_BIND 122 125 GTP (BY SIMILARITY).
FT DOMAIN 38 46 EFFECTOR REGION (PROBABLE).
FT LIPID 199 199 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 200 200 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 200 AA; 22755 MW; 2C658D153A290C30 CRC64;

Query Match 46.2%; Score 510; DB 1; Length 200;
Best Local Similarity 51.9%; Pred. No. 2.6e-34;
Matches 97; Conservative 39; Mismatches 43; Indels 8; Gaps 1;

QY 3 KOYDVLFRLLIGDSGVGKTCILCRFTDNEFHSHSTIGVDPKMTIEVDGIKVRIO 62
DB 4 KSYDYLKLLIGDSGVGKSCILLRSEDSFSPFITIGIDFKITFIELDGKRIKIQ 63
QY 63 DTAGQERYQITIKQYRRAGQIFLVYDISERSYQIHMKVSDVDVAPGVGKILGNK 122
DB 64 DTAGQERYFTITITAYYRGAMGIMLYDITNEKSFDRINRWIRNEHASADVEKMILGNK 123
QY 123 ADEQKRVQREGQQLAKEYGMDFYETSACTNLNKESFTRLTELVLQAHKREGLRM 182

```

```

184 ASNELALAELEEEEGKPEGPANSKTCWC 212
      | : : : | : : |
190 AQSPTIKPADQ-----SGNQAASACC 213

```

RESULT 13	ARA3 ARATH	STANDARD;	PRT;	216 AA.
UID	ARA3 ARATH			
ACC	P28186;			
DATE	01-DEC-1992	(Rel. 24, Created)		
DATE	01-DEC-1992	(Rel. 24, Last sequence update)		
DATE	15-JUN-2002	(Rel. 41, Last annotation update)		
DE	Kas-related protein ARA-3			
DE	ARA-3 OR AT3G46060 OR F12M12_30.			

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 NCBI_TaxID=3702.

strain-various strains; TISSUE=Leaf;
 MEDLINE=92084144; PubMed=1748311;
 Anai T., Hasegawa K., Watanabe Y., Uchimiya H., Ishizaki R.,
 Matsui M.;
 "Isolation and analysis of cDNAs encoding small GTP-binding proteins
 of *Arabidopsis thaliana*.";
 Gene 108:259-364(1991).

SEQUENCE FROM N.A.

X MEDLINE=21016720; PubMed=11130713;
 A SALANBANT M., LEMCKE K., RIEGER M., ANSORGE W., UNSELD M.,
 A FARTOUNCH B., VALLE G., BLOECKER H., PEREZ-MONSO M., OBERMAYER B.,
 A DELSNEY M., BOUTRY M., GRIVELL L.A., WACHE R., PUGDOMENAT P.,
 A DE SIMONE V., CHOISNE N., ARTIGUENAVE F., ROBERT C., BROTTIER P.,
 A WINCKER P., CATOLICO L., WEISSBACH J., SAURIN W., QUETIER F.,
 A SCHAERF M., MUELLER-AUER S., GIBBEL C.
 X STRAIN=CV. Columbia;

Wurbach E.; Drzonek H.; Erfle H.

Wiedelmann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G

vezzì A., D'Angelo M., Pallavicini A., Toppo S., Simionati B.,
Conrad A., Hornischer V., Vancini G.

A Reichelt J., Scharfe M., Schoen O.

Navarro P., Collado C., Perez-Perez A., Ottenwaelder B.

Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Masny D.

de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E., Monfort A.

Mannhaupt G., Haase D. School of

Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Zaccaria P., Mewes H.-W.,

Rooney T., Rizzo M., Walts A., Utterback T., Fuller C., V. Shea T. D.

Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
Pat C. Military & Civilian

Preuss D., Lin X, Nierman W C
Sellers P., Gill J.E., Feldblyum T.V.,

Fraser C.M., Kaneko T., Nakamura Y., Salzberg S.L., White O., Venter J.C.,

Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y., Yamamoto T., Sato S., Noto T., Asamizu E.,

Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.

Watanabe A., Nakazaki N., Shinpo S., Takeuchi C., Wada T., Nakayama S., Yamada M.

"Sequence and analysis of chromosome 2 of the

thaliana,";

Nature 408:820-822(2000).

----- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY. -----

This SWISS-PROT entry is converted to UniProt format.

between the Swiss Institute of Bioinformatics and the University of Geneva. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the University of Geneva. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the University of Geneva.

the European Bioinformatics Institute. There are no restrictions on

use by non-profit institutions as long as its content is in no

modified and this statement is not removed. Usage by entities requires a license agreement. Usage by and for commercial entities requires a license agreement.

or send an email to license@isb-sib.ch or send an email to license@isb-sib.ch (See <http://www.isb-sib.ch/announcements>)

----- (no. 30) -----

EMBL; D01025; BAA00830.1; -.

EMBL; AL333175; CAB90933.1; -.

```

DR PIR: J00908; J00908.
DR PIR: J00640; J00640.
DR HSP: P05713; 3RAB.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_trnsfrmg.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; ras; 1.
DR PRINTS: PR00449; RASTRNSFRMG.
DR SMART: SM00175; RAB; 1.
DR TIGRFAMs: TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Multigene family.
FT NP_BIND 22 29 GTP (BY SIMILARITY).
FT NP_BIND 70 74 GTP (BY SIMILARITY).
FT NP_BIND 128 131 GTP (BY SIMILARITY).
FT DOMAIN 44 52 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 213 213 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 214 214 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 216 AA; 23835 MW; 09E9C19A7A44E705 CRC64;

Query Match 46.0%; Score 508; DB 1; Length 216;
Best Local Similarity 46.9%; Pred. No. 4.2e-34;
Matches 100; Conservative 41; Mismatches 54; Indels 18; Gaps 3;

QY 5 YDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSHSTIGVDFKMTIEVDGKVRQIWD 64
DB 12 YDLKLLIGDSGVGKSCLLRFSGTSFTTIGDIFKRIEIDGGRKILQIWD 71

QY 65 AGQERYQITKQYRRAGGIFLVYDISSERSYQHIMKWSDVDEYAPGQVOKILGNKAD 124
DB 72 AGQERFRTITAYYRGAGILLVYDVDESFNIRWIRNIEHQASDNVNKILVGNKAD 131

QY 125 -EQRKRVGRGQQLAKEYGMDFYETSACTNLNKESFTLRLTELVLQAHKELEGLMR 183
DB 132 MDESKRAVPTAKGQALADEYKIFETTSKTNLNVVEEFSIG-----RDIQR 180

QY 184 ASNELALAE-----LEEKGKPEGPANSSKTC 210
DB 181 LSDTSRAEPATIKISQIDQAGAQATQKSAC 213

RESULT 14
YPT2_VOLCA
ID YPT2_VOLCA STANDARD; PRT; 217 AA.
AC P36861.
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE 01-JUN-1994 (Rel. 29, Last annotation update)
DE GTP-binding protein yptv2.
GN YPTV2.
OS Volvox carteri.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Volvocaceae; Volvox.
OX NCBI_TaxID=3067;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=f. Nagariensis / HK10;
RX MEDLINE=94037148; PubMed=8221932;
RA Fabry S., Jacobsen A., Huber H., Palme K., Schmitt R.;
RT "Structure, expression, and phylogenetic relationships of a family of
RT ypt genes encoding small G-proteins in the green alga Volvox
RT carteri."
RL Curr. Genet. 24:229-240(1993).
CC -!- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
CC TRAFFIC (BY SIMILARITY).
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL: L08128; AAA34251.1; -.
DR PIR: S36365; S36365.
DR HSP: P05713; 3RAB.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_trnsfrmg.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; ras; 1.
DR PRINTS: PR00449; RASTRNSFRMG.
DR SMART: SM00175; RAB; 1.
DR TIGRFAMs: TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport;
FT NP_BIND 20 27 GTP (BY SIMILARITY).
FT NP_BIND 68 72 GTP (BY SIMILARITY).
FT NP_BIND 126 129 GTP (BY SIMILARITY).
FT DOMAIN 42 50 EFFECTOR REGION (PROBABLE).
FT LIPID 215 215 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 216 216 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 217 AA; 24143 MW; 87D3B30C75689EAA CRC64;

Query Match 45.7%; Score 505.5; DB 1; Length 217;
Best Local Similarity 48.6%; Pred. No. 6.7e-34;
Matches 101; Conservative 40; Mismatches 64; Indels 3; Gaps 3;

QY 5 YDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSHSTIGVDFKMTIEVDGKVRQIWD 64
DB 10 YDLKLLIGDSGVGKSCLLRFTDNEFTTIGDIFKRIKVVGVGKLVKQIWD 69

QY 65 AGQERYQITKQYRRAGGIFLVYDISSERSYQHIMKWSDVDEYAPGQVOKILGNKAD 124
DB 70 AGQERFRTITAYYRGAGILLVYDISSERSYQHIMKWSDVDEYAPGQVOKILGNKAD 129

QY 125 -EQRKRVGRGQQLAKEYGMDFYETSACTNLNKESFTLRLTELVLQAHKELEGLMR 183
DB 130 LAEDKRVVSIARGQALADEFGFRFYETSAKDNVHVEEAFIAVKDVLARMEGEHANQQL 189

QY 184 ASNELALAE-LLEEKGKPEGPANSSKTC 210
DB 190 QQQQLSAAQPVRLTSGSP-SPAQKSCC 216

RESULT 15
RB10_RAT
ID RB10_RAT STANDARD; PRT; 200 AA.
AC P35281.
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Ras-related protein Rab-10.
GN RAB10.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=92210533; PubMed=1313420;
RA Eiferink L.A., Anzal K., Scheller R.H.;
RT "Rab15, a novel low molecular weight GTP-binding protein specifically
RT expressed in rat brain."
RL J. Biol. Chem. 267:5768-5775(1992).
CC -!- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND
CC NEUROTRANSMITTER RELEASE.
CC -!- TISSUE SPECIFICITY: HIGHEST LEVELS IN NEURAL AND MUSCLE TISSUES.
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

[illegible]

Search completed: June 18, 2003, 15:34:00
Job time : 23 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:25:06 ; Search time 81 Seconds
(without alignments)
539.284 Million cell updates/sec

Title: US-09-817-198b-2

Perfect score: 1105

Sequence: 1 MAKQYDVLFRLLIGDSGVG.....LEEEGKPEGPANSSKTCWC 212

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_virus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	798	72.2	168	11 Q91YW0	Q91YW0 mus musculus
2	557	50.4	211	5 Q9TYS2	Q9TYS2 caenorhabdi
3	546.5	49.5	204	5 O15971	O15971 drosophila
4	541	49.0	207	11 Q8VCF6	Q8VCF6 mus musculus
5	528.5	47.8	207	5 O18338	O18338 drosophila
6	526.5	47.6	200	4 Q9H0T3	Q9H0T3 homo sapien
7	526	47.6	202	11 Q9DD03	Q9DD03 mus musculus
8	521	47.1	206	3 Q9HET4	Q9HET4 aspergillus
9	519	47.0	216	10 Q24466	Q24466 arabidopsis
10	518	46.9	214	10 Q40218	Q40218 lotus japon
11	515.5	46.7	216	10 Q40215	Q40215 lotus japon
12	515	46.6	216	10 Q9FJF1	Q9FJF1 arabidopsis
13	514	46.5	201	5 Q94148	Q94148 caenorhabdi
14	513.5	46.5	215	10 Q40219	Q40219 lotus japon
15	513	46.4	216	10 Q8VWF9	Q8VWF9 nicotiana t
16	512	46.3	216	10 Q40177	Q40177 lycopersico

17 512 46.3 216 10 Q9LZD4
18 511.5 46.3 212 10 Q40217
19 510.5 46.2 212 10 Q8W3J4
20 510 46.2 218 10 Q9SF91
21 508.5 46.0 215 10 Q41023
22 508 46.0 216 10 Q49844
23 508 46.0 216 10 Q8W3J3
24 507.5 45.9 222 10 Q96362
25 507 45.8 202 3 Q96VL3
26 506.5 45.8 216 10 Q9SWV8
27 506 45.8 216 10 Q41024
28 505 45.7 198 4 Q96GU4
29 505 45.5 215 10 Q8W3J2
30 502.5 44.8 204 3 Q9C1X5
31 495.5 44.8 204 3 Q9C1X5
32 492.5 44.6 215 10 Q41061
33 484 43.8 205 5 Q9BLE3
34 482 43.6 203 10 Q40569
35 482 43.6 205 5 Q26554
36 481 43.5 203 10 Q8RU63
37 478.5 43.3 201 4 Q9H0U4
38 478 43.3 202 10 Q9ZRH6
39 477 43.2 202 3 Q9HDT5
40 477 43.2 202 10 Q39845
41 477 43.2 202 10 Q8L155
42 476 43.1 203 10 Q94027
43 474 42.9 205 10 Q9ZRE2
44 473.5 42.9 201 11 Q9DJG1
45 473.5 42.9 205 5 Q9UAQ6

ALIGNMENTS

RESULT 1

Q91YW0 ID Q91YW0 PRELIMINARY; PRT; 168 AA.
AC Q91YW0;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 19.4 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Cranial; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC013790; AAH13790.1; -
DR InterPro; IPR001806; Ras_trnsmfing.
DR InterPro; IPR002078; Sig54_interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR TIGRFAMs; TIGR00231; Small_GTP; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW GTP-binding; Hypothetical protein.
SQ SEQUENCE 168 AA; 19399 MW; 4E2C2FB1C56BCDF8 CRC64;

Query Match 72.2%; Score 798; DB 11; Length 168;
Best Local Similarity 99.3%; Pred. No. 3.7e-59;
Matches 152; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMTIEVDGKVRQ 60
DB 1 MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFKMTIDVDGKVRQ 60
QY 61 IWDTAGQERYQTITKQYVRRAGQIFLVYDIDISSERSYQHIMKWSVDVEYAPGVOKILIG 120
DB 61 IWDTAGQERYQTITKQYVRRAGQIFLVYDIDISSERSYQHIMKWSVDVEYAPGVOKILIG 120
QY 121 NKADEQKRVQREGGQQLAKEYGMDFYTSAC 153

DR TIGR00231; small_GTP; 1.
 DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
 KW GTP-binding; Lipoprotein.
 SQ SEQUENCE 204 AA; 23336 MW; 4E058761C6854920 CRC64;

Query Match 49.5%; Score 546.5; DB 5; Length 204;
 Best Local Similarity 49.8%; Pred. No. 5e-38;
 Matches 105; Conservative 47; Mismatches 50; Indels 9; Gaps 3;
 Qy 1 MAKQYDVLRLLIGDSGVGKTCCLCRFTDNEFHSHSTIGYDFKMTIEVDGIKVR 59
 Db 1 MAKYDVLRLKLLIGDSGVGKTCCLFRSDDAFTSTFSTIGIDFKIKTVLRGKKIL 60
 Qy 60 QIWDTAGQERYQTITKQYRRAQGIPLVYDISSERSYQHIMKWSDVDEYAPEGVQKILI 119
 Db 61 QIWDTAGQERFHTITTYIRGAMGIMLVYDITNEKSFENYKWLNRINDEHANEDVERKIL 120
 Qy 120 GNKADEQKRVGQGQQLAKKEYGMDYFETSACTNLNIKESFTRLTVELVLOAHKELEG 179
 Db 121 GNKCDMTDKRVNKEGEATAREHGIRFEMETSAKSNINIERAFCELAAILD----KTSG 176
 Qy 180 LMRASNELALAELEEEGPEGPANSSKTC 210
 Db 177 RESAENQERVLIIDRRNQEKAP----GYSKCC 203

RESULT 4

Q8VCF6 PRELIMINARY; PRT; 207 AA.
 AC Q8VCF6;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Similar to mel transforming oncogene (derived from cell line
 DE NK14)-RAB8 homolog.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Strausberg R;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC019930; AHH19930.1; -
 DR InterPro; IPR003579; GTPase_Rab.
 DR InterPro; IPR001230; Prenyl_site.
 DR InterPro; IPR001806; Ras_trnsfmg.
 DR InterPro; IPR002078; Sig54_interact.
 DR Pfam; PF00071; ras; 1.
 DR PRINTS; PRO0449; RASTRNSFRMG.
 DR SMART; SM00175; Rab; 1.
 DR PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
 DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
 SQ SEQUENCE 207 AA; 23668 MW; AC89DC85588FB8F8 CRC64;

Query Match 49.08%; Score 541; DB 11; Length 207;
 Best Local Similarity 50.3%; Pred. No. 1.5e-37;
 Matches 99; Conservative 49; Mismatches 49; Indels 0; Gaps 0;
 Qy 1 MAKQYDVLRLLIGDSGVGKTCCLCRFTDNEFHSHSTIGYDFKMTIEVDGIKVR 60
 Db 1 MAKYDVLRLKLLIGDSGVGKTCCLFRSDDAFTSTFSTIGIDFKIKTVLRGKKILQ 60
 Qy 61 IWDTAGQERYQTITKQYRRAQGIPLVYDISSERSYQHIMKWSDVDEYAPEGVQKILIG 120
 Db 61 IWDTAGQERFHTITTYIRGAMGIMLVYDITNEKSFENYKWLNRINDEHANEDVERKILG 120
 Qy 121 NKAEDEQKRVGQGQQLAKKEYGMDYFETSACTNLNIKESFTRLTVELVLOAHKELEG 180
 Db 121 NKCDVNDKRVSKERGEKALDYGIKPFMETSAKININVENAFTFLARDIKAKMKKEGN 180
 Qy 181 RMRASNELALAELEEE 197

Db 181 SPOGSSHGKTVTEQOK 197

RESULT 5

O18338 PRELIMINARY; PRT; 207 AA.
 AC O18338;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE RAB8 protein (LD44762p).
 GN RAB8 OR CG8287.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RC MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OREGON R; TISSUE=HEAD;
 RC MEDLINE=97228579; PubMed=9074639;
 RA Satoh A.K., Tokunaga F., Ozaki K.;
 RA "Rab proteins of Drosophila melanogaster: novel members of the Rab-
 RT protein family.";
 RL FEBS Lett. 404:65-69(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,

DR HSP: P05713; 3RAB.
 DR MGD; MGI.191578; 0610007N03rik.
 DR InterPro: IPR003579; GTPase_Rab.
 DR InterPro: IPR001230; Prenyl_site.
 DR InterPro: IPR001806; Ras_trnsmg.
 DR InterPro: IPR002078; Sig54_interact.
 DR InterPro: IPR005225; Small_GTP.
 DR Pfam: PF00071; ras; 1.
 DR PRINTS: PR00449; RASTRNSFRMG.
 DR SMART: SM00175; RAB; 1.
 DR TIGRFAMS: TIGR00231; small_gtp; 1.
 DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.
 DR PROSITE: PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
 KW GTP-binding; Lipoprotein.
 SQ SEQUENCE 202 AA; 22770 MW; 5DF599432E228AC0 CRC64;
 Query Match 47.6%; Score 526; DB 11; Length 202;
 Best Local Similarity 54.1%; Pred. No. 2.5e-36;
 Matches 92; Conservative 45; Mismatches 33; Indels 0; Gaps 0;
 QY 1 MAKQYDLFRLLIGDSGVGKTCLLCRFTDNEFHSHSHISTIGVDYFKMTIEVDGKVRQ 60
 DB 1 MAKAYDLFRLLIGDSGVGKTCLLIRFAEDNFNSTYISTIGIDFKIRTVDIEGKRIK 60
 QY 61 IWDAGERYQYITKQYRRAQGIIFYVDISSERSYOHIMKWVSDVDYAPEGVQKILIG 120
 DB 61 VMDTAGERKTIITAYIRGAGIILYDITDEKSFENIQNMWKSINENASAGVERLLG 120
 QY 121 NKADEQKRVGREGQQLAKYGMDFYETSACTNLNIKESFTRLTSLVL 170
 DB 121 NKCDMEAKRVQREAEKLAREHRIREFEISAKSSVNVDEAFSSLDAIL 170
 RESULT 8
 Q9HET4 PRELIMINARY; PRT; 206 AA.
 AC Q9HET4;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Secretion related GTPase, (SrgA).
 GN SrgA.
 OS Aspergillus niger.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5061;
 RN [1]
 RC STRAIN=N402;
 RA Punt P.J., Seiboth B., Weenink X.O., van Zeijl C.M., Lenders M.,
 RA Van Den Hondel C.A.;
 RA Konetschny C., Ram A.F., Montijn R., Kubicek C.P.,
 RA "Identification and characterisation of a family of secretion related
 RT small GTPase encoding genes from the filamentous fungus Aspergillus
 RT niger; a putative SEC4 homologue is not essential for growth.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 DR EMBL: AJ278658; CAC17832.1; -.
 DR HSP: P05713; 3RAB.
 DR InterPro: IPR003579; GTPase_Rab.
 DR InterPro: IPR001806; Ras_trnsmg.
 DR InterPro: IPR005225; Small_GTP.
 DR Pfam: PF00071; ras; 1.
 DR PRINTS: PR00449; RASTRNSFRMG.
 DR SMART: SM00175; RAB; 1.
 DR TIGRFAMS: TIGR00231; small_gtp; 1.
 KW GTP-binding; Lipoprotein.
 SQ SEQUENCE 206 AA; 22823 MW; 501916B795CF8C8C CRC64;
 Query Match 47.1%; Score 521; DB 3; Length 206;
 Best Local Similarity 60.6%; Pred. No. 6.8e-36;
 Matches 97; Conservative 30; Mismatches 33; Indels 0; Gaps 0;

QY 3 KOYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHSHISTIGVDYFKMTIEVDGKVRQIW 62
 DB 5 RNYDFLIKLLIGDSGVGKSCCLLRFSDSFTPTFTTIGIDFKIRTIELDKGRVQLQIW 64
 QY 63 DTAGERYQYITKQYRRAQGIIFYVDISSERSYOHIMKWVSDVDYAPEGVQKILIGNK 122
 DB 65 DTAGQERFRITAYIRGAGIILYDITDERSFQNTWFSNVEQHASQGVHKLIGNK 124
 QY 123 ADEQKRVGREGQQLAKYGMDFYETSACTNLNIKESF 162
 DB 125 CDWEKRAVSTEQQLANELGIPFLEVSANKNNINIEKAF 164
 RESULT 9
 O24466 PRELIMINARY; PRT; 216 AA.
 AC O24466;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE ATRAB8 (GTPase ATRAB8).
 GN RAB8 OR F4P12_310.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Bischoff F., Palme K.;
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RC SEQUENCE FROM N.A.
 RA Bloecker H., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X., Quetier F.,
 RA Salancoubat M.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RC SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 DR EMBL: U82434; AAB65088.1; -.
 DR EMBL: AL132966; CAB67668.1; -.
 DR HSP: P05713; 3RAB.
 DR InterPro: IPR003579; GTPase_Rab.
 DR InterPro: IPR001806; Ras_trnsmg.
 DR InterPro: IPR002078; Sig54_interact.
 DR InterPro: IPR005225; Small_GTP.
 DR Pfam: PF00071; ras; 1.
 DR PRINTS: PR00449; RASTRNSFRMG.
 DR SMART: SM00175; RAB; 1.
 DR TIGRFAMS: TIGR00231; small_gtp; 1.
 DR PROSITE: PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
 KW GTP-binding; Lipoprotein.
 SQ SEQUENCE 216 AA; 23939 MW; 1648E45B29D4EEB6 CRC64;
 Query Match 47.0%; Score 519; DB 10; Length 216;
 Best Local Similarity 51.9%; Pred. No. 1.1e-35;
 Matches 98; Conservative 42; Mismatches 43; Indels 6; Gaps 2;
 QY 5 YDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHSHISTIGVDYFKMTIEVDGKVRQIWDT 64
 DB 12 YDYLKLLIGDSGVGKSCCLLRFSDSFTPTFTTIGIDFKIRTIELDKGRKRIQWDT 71
 QY 65 AGERYQYITKQYRRAQGIIFYVDISSERSYOHIMKWVSDVDYAPEGVQKILIGNKAD 124
 DB 72 AGQERFRITAYIRGAGIILYDITDESFNIRNIRNIEQHASQGVHKLIGNKAD 131
 QY 125 -EQKRVGREGQQLAKYGMDFYETSACTNLNIKESFTRLTSLVLA-----HRKELE 178
 DB 132 MDESKRAVPKSKQALADEYGMKFEFESAKTNLNVVEFFSIADIKRIQADTARAEQ 191

CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 DR EMBL; AB015475; BAB08351.1; -

DR HSP; P05713; 3RAB.
 DR InterPro; IPR003579; GTPase_Rab.
 DR InterPro; IPR001806; Ras_trnsmg.
 DR InterPro; IPR002078; Sig54_interact.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00071; ras; 1.
 DR PRINTS; PR00449; RASTRNSFRMG.
 DR SMART; SM00175; RAB; 1.
 DR TIGRFAMS; TIGR00231; small_gtp; 1.
 DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
 KW GTP-binding; Lipoprotein.
 SQ SEQUENCE 216 AA; 23834 MW; E1370672F6E3364 CRC64;

Query Match 46.6%; Score 515; DB 10; Length 216;
 Best Local Similarity 47.6%; Pred. No. 2.3e-35;
 Matches 101; Conservative 43; Mismatches 52; Indels 16; Gaps 3;

QY 5 YDVLFRLLIGDSGVGKTCILCRFTDNEFHSSHSHTIGVDFKMTIEVDGKIVRIQIWD 64
 DB 12 YDYLKLLIGDSGVGKSCLLRFSFGSFTTSFTTIGIDFKINTIELDKRIKLIQIWD 71
 QY 65 AGERYOTITKQYVYRAQGIYVYDISSERSYOHIMKWSDVDYAPGEGVOKILGNKAD 124
 DB 72 AGQERFRTITAYRGAMGILLVYDVDESFNIRNIRNIEGHASDNVKNILVGNKAD 131
 QY 125 -EQKRGVQEQOQAKKEYGMDFYETSACTNLNLIKESFTRLTLVLQA-----HRKELE 178
 DB 132 MDESKRAVPSKQALADEYGIKFETSAKTNLNVVEVFFSIADIKQRLADTDSRAEPA 191
 QY 179 GLMRASNELALAELEEKGPEGPANSKTC 210
 DB 192 TIKISQDQAAG-----GQATQKSAC 213

RESULT 13

Q94148 PRELIMINARY; PRT; 201 AA.
 AC Q94148;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE RAB10-like.
 GN T3H2.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Wansley P., Bradshaw H.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-124 FROM N.A.
 RA Nonet M.E., Staunton J.E., Kilgard M.P., Fergestad T., Jorgensen E.,
 RA Hartweig E., Horvitz H.R., Meyer B.J.;
 RA Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 DR EMBL; U80033; AAC48200.1; -
 DR EMBL; U68257; AAB16972.1; -
 DR HSP; P05713; 3RAB.
 DR InterPro; IPR003579; GTPase_Rab.
 DR InterPro; IPR001806; Ras_trnsmg.
 DR InterPro; IPR002078; Sig54_interact.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00071; ras; 1.
 DR PRINTS; PR00449; RASTRNSFRMG.
 DR SMART; SM00175; RAB; 1.
 DR TIGRFAMS; TIGR00231; small_gtp; 1.
 DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
 KW GTP-binding; Lipoprotein.

SQ SEQUENCE 201 AA; 22712 MW; 2D205ABF751EBF1A CRC64;

Query Match 46.5%; Score 514; DB 5; Length 201;
 Best Local Similarity 46.9%; Pred. No. 2.5e-35;
 Matches 100; Conservative 48; Mismatches 51; Indels 14; Gaps 3;
 QY 1 MAKQ-YDVLFRLLIGDSGVGKTCILCRFTDNEFHSSHSHTIGVDFKMTIEVDGKIVRI 59
 DB 1 MARRPYDMLFKLLIGDSGVGKTCILYRFSDDAFTTISTIGIDFKIKTIELKGGKIKL 60
 QY 60 QIWDTAGQERYOTITKQYVYRAQGIYVYDISSERSYOHIMKWSDVDYAPGEGVOKILI 119
 DB 61 QIWDTAGQERFRTITAYRGAMGILLVYDITNAKSDNTAKWIRNDEHASEVDVVKIL 120
 QY 120 GNKADEPKRGVQEQOQAKKEYGMDFYETSACTNLNLIKESFTRLTLVLQAHRKELEG 179
 DB 121 GNKCDMSDRVRSERGEKIAQDHGISFHTSAKLNVDVAFYDLAEAIL----- 171
 QY 180 LMRASNELALAELEEKGPEGPANSKTCWC 212
 DB 172 AKMPDSTD-----EQSRDTVNPVQPORQSSGGC 200

RESULT 14

Q40219 PRELIMINARY; PRT; 215 AA.
 AC Q40219;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE RAB8E.
 GN RAB8E.
 OS Lotus japonicus.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.
 OX NCBI_TaxID=34305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=ROOT NODULES;
 RX MEDLINE=97231679; PubMed=9076991;
 RA Borg S., Brandstrup B., Jensen T.J., Poulsen C.;
 RT "Identification of new protein species among 33 different small GTP-binding proteins encoded by cDNAs from Lotus japonicus, and expression of corresponding mRNAs in developing root nodules."
 RL Plant J. 11:237-250(1997).
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 DR EMBL; Z73948; CAA98176.1; -
 DR HSP; P05713; 3RAB.
 DR InterPro; IPR003579; GTPase_Rab.
 DR InterPro; IPR001806; Ras_trnsmg.
 DR InterPro; IPR002078; Sig54_interact.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00071; ras; 1.
 DR PRINTS; PR00449; RASTRNSFRMG.
 DR SMART; SM00175; RAB; 1.
 DR TIGRFAMS; TIGR00231; small_gtp; 1.
 DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
 KW GTP-binding; Lipoprotein.
 SQ SEQUENCE 215 AA; 23715 MW; E7912846E919D608 CRC64;

Query Match 46.5%; Score 513.5; DB 10; Length 215;
 Best Local Similarity 46.9%; Pred. No. 3.1e-35;
 Matches 100; Conservative 43; Mismatches 53; Indels 17; Gaps 3;

QY 5 YDVLFRLLIGDSGVGKTCILCRFTDNEFHSSHSHTIGVDFKMTIEVDGKIVRIQIWD 64
 DB 12 YDYLKLLIGDSGVGKSCLLRFSFGSFTTSFTTIGIDFKIRTIELDKRIKLIQIWD 71
 QY 65 AGERYOTITKQYVYRAQGIYVYDISSERSYOHIMKWSDVDYAPGEGVOKILGNKAD 124
 DB 72 AGQERFRTITAYRGAMGILLVYDVDESFNIRNIRNIEGHASDNVKNILVGNKAD 131

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:32:26 ; Search time 27 Seconds
(without alignments)
231.024 Million cell updates/sec

Title: US-09-817-198B-2
Perfect score: 1105
Sequence: 1 MAKQYDVLFRLLIGDSVG.....LEEKGKPEGPANSSKTCWC 212

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs; 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCFUS-COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	545	49.3	207	2	US-08-824-873-4
2	545	49.3	207	3	Sequence 4, Appli
3	527.5	47.7	205	2	US-09-198-184-4
4	527.5	47.7	205	2	Sequence 25, Appl
5	518.5	46.9	198	2	US-08-718-270A-25
6	518.5	46.9	198	2	Sequence 51, Appl
7	503.5	45.6	207	2	US-08-531-525-51
8	503.5	45.6	207	2	Sequence 35, Appl
9	494.5	44.8	215	2	US-08-718-270A-35
10	494.5	44.8	215	2	Sequence 10, Appl
11	486.5	44.2	194	2	US-08-531-525-10
12	486.5	44.2	194	2	Sequence 34, Appl
13	478.5	43.3	201	2	US-08-718-270A-34
14	478.5	43.3	201	2	Sequence 3, Appli
15	469	42.4	201	2	US-09-154-602-3
16	469	42.4	201	2	Sequence 8, Appli
17	463	41.9	202	2	US-08-916-901-8
18	463	41.9	202	2	Sequence 14, Appl
19	453.5	41.0	201	2	US-08-531-525-14
20	453.5	41.0	201	2	Sequence 13, Appl
21	452.5	41.0	190	2	US-08-718-270A-13
22	452.5	41.0	190	2	Sequence 3, Appli
23	443	40.1	218	2	US-08-824-873-3
24	443	40.1	218	2	Sequence 3, Appli
25	416	37.6	191	4	US-08-531-525-19
26	412.5	37.3	212	4	US-08-718-270A-19
27	409.5	37.1	190	2	US-09-075-454-3
					Sequence 67, Appl
					Sequence 1, Appli

28	409.5	37.1	190	3	US-09-198-184-1	Sequence 1, Appli
29	409.5	37.1	212	2	US-08-531-525-18	Sequence 18, Appl
30	409.5	37.1	212	2	US-08-718-270A-18	Sequence 18, Appl
31	409.5	37.1	214	2	US-08-531-525-52	Sequence 52, Appl
32	409.5	37.1	214	2	US-08-718-270A-52	Sequence 52, Appl
33	396.5	35.9	203	2	US-08-766-551-8	Sequence 8, Appli
34	387	35.0	210	2	US-08-531-525-16	Sequence 16, Appl
35	387	35.0	210	2	US-08-718-270A-16	Sequence 16, Appl
36	373	33.8	213	2	US-08-773-423-8	Sequence 8, Appli
37	373	33.8	217	2	US-08-773-423-3	Sequence 3, Appli
38	372	33.7	208	2	US-08-531-525-17	Sequence 17, Appl
39	372	33.7	208	2	US-08-718-270A-17	Sequence 17, Appl
40	368.5	33.3	208	2	US-08-531-525-15	Sequence 15, Appl
41	368.5	33.3	208	2	US-08-718-270A-15	Sequence 15, Appl
42	360	32.6	213	2	US-08-531-525-11	Sequence 11, Appl
43	360	32.6	213	2	US-08-718-270A-11	Sequence 11, Appl
44	359.5	32.5	191	2	US-08-531-525-26	Sequence 26, Appl
45	359.5	32.5	191	2	US-08-718-270A-26	Sequence 26, Appl

ALIGNMENTS

RESULT 1
US-08-824-873-4
; Sequence 4, Application US/08824873
; Patent No. 5843717
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Guegler, Karl
; TITLE OF INVENTION: NOVEL RAB PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/824,873
; FILING DATE: Filed Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0240 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 207 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 234746
US-08-824-873-4

Query Match 49.3%; Score 545; DB 2; Length 207;
Best Local Similarity 52.9%; Pred. No. 2.2e-52;
Matches 99; Conservative 45; Mismatches 43; Indels 0; Gaps 0;

Db 121 CDVNDKQVSKERGEKALDYGIKFMETSAKANINVENAFTLARDIKAKMDKKLEGNSP 180
Qy 183 RASNE 187
Db 181 QGSNQ 185

RESULT 4

US-08-718-270A-25
; Sequence 25, Application US/08718270A
; Patent No. 5910478
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 5910478le, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptidomimetics Inhibiting
; TITLE OF INVENTION: the Oncogenic Action of P21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,270A
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/531,525
; FILING DATE: 21-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,091
; FILING DATE: 21-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 78-95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8089
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 205 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Canis familiaris
US-08-718-270A-25

Query Match 47.7%; Score 527.5; DB 2; Length 205;
Best Local Similarity 52.4%; Pred. No. 1.9e-50;
Matches 97; Conservative 45; Mismatches 42; Indels 1; Gaps 1;
Qy 3 KQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSHISTIGVDFKMKTIKIEVDGKVRQIW 62
Db 2 KTYDLFKLLIGDSGVGKTCVLFRESDAFNFTISIGIDFKITVELQKKIKLQIW 61
Qy 63 DTAGQERYQTITKQYRRAGQIFLVYDISSERSYOHIMKWVSDVDEYAPGEGVOKILGNK 122
Db 62 DTAGQERYTITAYY-RANGIMLVYDITNEKSPDNRNIRNIEERHASADVEKWLGNK 120

Qy 123 ADEBQKQVREGQQQALAKYGMDFYETSACTNNIKESFTRLRLVLAHRKLEGLRM 182
Db 121 CDVNDKQVSKERGEKALDYGIKFMETSAKANINVENAFTLARDIKAKMDKKLEGNSP 180
Qy 183 RASNE 187
Db 181 QGSNQ 185
RESULT 5
US-08-531-525-51
; Sequence 51, Application US/08531525
; Patent No. 5840683
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 5840683le, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action
; TITLE OF INVENTION: of P21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531,525
; FILING DATE: 21-SEP-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 37-94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8089
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Canis familiaris
US-08-531-525-51

Query Match 46.9%; Score 518.5; DB 2; Length 198;
Best Local Similarity 55.0%; Pred. No. 1.8e-49;
Matches 94; Conservative 45; Mismatches 31; Indels 1; Gaps 1;
Qy 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSHISTIGVDFKMKTIKIEVDGKVRQIW 60
Db 1 MKTYDLFKLLIGDSGVGKTCVLFRESDAFNFTIS-IGIDFKITVELQKKIKLQ 59
Qy 61 IWDTAGQERYQTITKQYRRAGQIFLVYDISSERSYOHIMKWVSDVDEYAPGEGVOKILIG 120
Db 60 IWDTAGQERYTITTSYRGAMGLMVLVDITNGKSFENISKWLNRIDEHANEDVERMLG 119
Qy 121 NKADEEQKQVREGQQQALAKYGMDFYETSACTNNIKESFTRLRLVLA 171
Db 120 NKCDMDKRVVPVKGQIAREHGIRFFETSAKVYNINIERAFLTLAEDILR 170


```

; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 59104781e, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptidomimetics Inhibiting
; the Oncogenic Action of P21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,270A
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/531,525
; FILING DATE: 21-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,091
; FILING DATE: 21-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 78-95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8089
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 207 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Discopyge ommata
; US-08-718-270A-35

```

```

Query Match 45.68; Score 503.5; DB 2; Length 207;
Best Local Similarity 46.4%; Pred. No. 8.8e-46;
Matches 96; Conservative 56; Mismatches 44; Indels 11; Gaps 4;

Qy 3 KOYDVLPRLLIGDSGVGKTCLLCRFTDNEFHSHSHSTIGVDFKMTIEVDGKVRQIWI 62
Db 2 KTYDLPKLLIGDSGVGKTCLLCRFTDNEFHSHSHSTIGVDFKMTIEVDGKVRQIWI 61

Qy 63 DTAGQERYQITKQYRRAGIFLVYDYSERVOHMKWSDVDEYAPGVOKILGNK 122
Db 62 DTAGQERFRTIT-AYYGAMGIMKV-DITNEKSPDNKINIRNIEEHASSDVERMILGNK 119

Qy 123 ADEQKRVGREGQQQLAKYGMDFYETSACTNLNIKESFTRLTETVLQAHKKEGLELRM 182
Db 120 CDMEKRVQKREGKEKLAIDYGIKFLKTSKSSINVEEAFITLARDINTLKNKK- 174

Qy 183 RASNELALAE-LEEEGKPEGPANSSK 208
Db 175 --NENSLQEAVDKLSPPKPKSOKK 198

```

```

RESULT 9
US-08-531-525-10
; Sequence 10, Application US/08531525

```

```

; Patent No. 5840683
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 58406831e, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action
; the Oncogenic Action of P21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531,525
; FILING DATE: 21-SEP-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 37-94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8089
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Arabidopsis thaliana
; US-08-531-525-10

```

```

Query Match 44.88; Score 494.5; DB 2; Length 215;
Best Local Similarity 46.9%; Pred. No. 9.2e-47;
Matches 100; Conservative 40; Mismatches 54; Indels 19; Gaps 4;

Qy 5 YDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHSHSTIGVDFKMTIEVDGKVRQIWI 64
Db 12 YDYLKLLIGDSGVGKTCLLCRFTDNEFHSHSHSTIGVDFKMTIEVDGKVRQIWI 71

Qy 65 AGQERYQITKQYRRAGIFLVYDYSERVOHMKWSDVDEYAPGVOKILGNKAD 124
Db 72 AGQER-RTITAYRGAMGILLVYDVTDESFNIRNIRNIEEHASSDVERMILGNKAD 130

Qy 125 -EOKRVGREGQQQLAKYGMDFYETSACTNLNIKESFTRLTETVLQAHKKEGLELRM 183
Db 131 MDESKRAVPTAKQALADEYGIKFFETSAKTNLNVVEVFFSIG- 179

Qy 184 ASNELALAE-----LEEEGKPEGPANSSKTC 210
Db 180 LSDTDSRAEPATIKISQTDQAAGAQATQKSAC 212

```

```

RESULT 10
US-08-718-270A-10
; Sequence 10, Application US/08718270A
; Patent No. 5910478
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.

```

APPLICANT: No. 59104781e, John F.
APPLICANT: Abajian, Henry B.
APPLICANT: Kende, Andrew S.
TITLE OF INVENTION: Peptidomimetics Inhibiting
the Oncogenic Action of P21 Ras
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/718,270A
FILING DATE: 20-SEP-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/531,525
FILING DATE: 21-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,091
FILING DATE: 21-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 78-95
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 215 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Arabidopsis thaliana
US-08-718-270A-10

Query Match 44.8%; Score 494.5; DB 2; Length 215;
Best Local Similarity 46.9%; Pred. No. 9,2e-47;
Matches 100; Conservative 40; Mismatches 54; Indels 19; Gaps 4;
QY 5 YDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHTIGVDGKVKRIQWDT 64
DB 12 YDLKLLIGDSGVGKSCLLRFSGTSFTTIGIDFKIRTIELDKRIKLIQWDT 71
QY 65 AQERYQITKQYRRAGQIFLYVDISSERSYQHIMKWSDVDYAPGVQKILGNAD 124
DB 72 AQGER-RITATTAYRGAMGILLYVDVTDESSFNIRNIRIEQHASDNVKNILVGNAD 130
QY 125 -EEQKRVGREGQQLAKKEYGMDFYETSACTNLNLIKESFTRTELVLQAHRELGLRMR 183
DB 131 MDESRAVPTAKQALADEYGIKFETSAKTNLNEVEVFFSIG-----RDIKOR 179
QY 184 ASNELALAE-----LEEKEGPEGPANSSKTC 210
DB 180 LSDTDSRAEPATIKISQTDQAAGAGQATQKSAC 212

RESULT 11
US-08-531-525-34
Sequence 34, Application US/08531525
Patent No. 5840683
GENERAL INFORMATION:

APPLICANT: Hlavka, Joseph J.
APPLICANT: Pincus, Matthew R.
APPLICANT: No. 58406831e, John F.
APPLICANT: Abajian, Henry B.
APPLICANT: Kende, Andrew S.
TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action
the Oncogenic Action of P21 Ras
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee and Winner, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/531,525
FILING DATE: 21-SEP-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 37-94
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 194 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Discopyge ommata
US-08-531-525-34

Query Match 44.2%; Score 488.5; DB 2; Length 194;
Best Local Similarity 55.8%; Pred. No. 3.6e-46;
Matches 96; Conservative 41; Mismatches 30; Indels 5; Gaps 4;
QY 1 MAKQ-YDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHTIGVDGKVKRIQWDT 59
DB 1 MAKYDILFRLLIGDSGVGKTCVLFPSDDAFNITISTIGIDFKINTVELGKKIKL 60
QY 60 QIWDTAGQERYQITKQYRRAGQIFLYVDISSERSYQHIMKWSDVDYAPGVQKILI 119
DB 61 QIWDTAGQERYQIT-SYIRGAMGIMLYVDITNAKSFENISKWLNIDEHANEVDYRMLL 119
QY 120 GNKADEQRQVGRGQQLAKKEYGMDFYETSACTNLNLIKESFTRTELVLQ 171
DB 120 GNK-DMEDKRVYLVKSGQ--IAEHAIRFFETSAKANINIERAFLTLAEDILQ 168

RESULT 12
US-08-718-270A-34
Sequence 34, Application US/08718270A
Patent No. 5910478
GENERAL INFORMATION:
APPLICANT: Hlavka, Joseph J.
APPLICANT: Pincus, Matthew R.
APPLICANT: No. 59104781e, John F.
APPLICANT: Abajian, Henry B.
APPLICANT: Kende, Andrew S.
TITLE OF INVENTION: Peptidomimetics Inhibiting
the Oncogenic Action of P21 Ras
NUMBER OF SEQUENCES: 52

Query Match	43.3%	Score 478.5;	DB 2;	Length 201;
Best Local Similarity	45.0%	Pred. No. 4,9e-45;		

[illegible]

```

RESULT 14
US-09-154-602-3
Sequence 3, Application US/09154602
Patent No. 6300472
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: RAB PROTEIN
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette

```

```
;
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/154,602
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/916,901
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0367 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: LIVRTUT04
; CLONE: 2514506
; US-09-154-602-3

Query Match 43.3%; Score 478.5; DB 4; Length 201;
Best Local Similarity 45.0%; Pred. No. 4.9e-45;
Matches 91; Conservative 42; Mismatches 64; Indels 5; Gaps 1;

QY 1 MAKQYDVLFRLLIGDSCVGKTCCLRCFTDNEFHSSHSITIGVDPKMTIEVDGKVRIG 60
Db 1 MPPEYDYLKLLIGDSCVGKSCLLRFPADDTYTESYSTIGVDPKRTIELDGKTIKLG 60
QY 61 IWDTAGQERYQITIKQYRRAGGIFLVYDIDSSERSYQHIMKWSVDVEYAPGKILIG 120
Db 61 IWDTAGQERFRTITSSYYRGAGHIIYVDVTDQESYANVKWLQEIYASENVNKLVG 120
QY 121 NKADEEQKQVREGQOQLAKKEYGMDFYETSACTNLNKESTRLTELVLQHRKEBGL 180
Db 121 NKSDLTTRKVDNTAKEFADSLGIPFLETSAKNATNVEQAF-----MTMAAEIKRMGP 175
QY 181 RMRASNELALAELEEKGPEG 202
Db 176 GAASGGERPNLKIDSTPVKPG 197

RESULT 15
US-08-916-901-8
; Sequence 8, Application US/08916901
; Patent No. 5892012
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: RAB PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/916,901

;
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0367 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 57006
; US-08-916-901-8

Query Match 42.4%; Score 469; DB 2; Length 201;
Best Local Similarity 51.2%; Pred. No. 5.5e-44;
Matches 83; Conservative 36; Mismatches 43; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDSCVGKTCCLRCFTDNEFHSSHSITIGVDPKMTIEVDGKVRIG 60
Db 1 MPPEYDYLKLLIGDSCVGKSCLLRFPADDTYTESYSTIGVDPKRTIELDGKTIKLG 60
QY 61 IWDTAGQERYQITIKQYRRAGGIFLVYDIDSSERSYQHIMKWSVDVEYAPGKILIG 120
Db 61 IWDTAGQERFRTITSSYYRGAGHIIYVDVTDQESYANVKWLQEIYASENVNKLVG 120
QY 121 NKADEEQKQVREGQOQLAKKEYGMDFYETSACTNLNKESTRLTELVLQHRKEBGL 162
Db 121 NKSDLTTRKVDNTAKEFADSLGIPFLETSAKNATNVEQAF 162

Search completed: June 18, 2003, 15:36:59
Job time : 37 secs
```


RESULT 2

```
US-09-764-868-701
; Sequence 701, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 701
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-868-701

Query Match      100.0%; Score 1105; DB 9; Length 401;
Best Local Similarity 100.0%; Pred. No. 8e-96;
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDSVGKTCCLCRFTDNEFHSHSHISTIGVDFKMTIEVDGKVRQ 60
DB 27 MAKQYDVLFRLLIGDSVGKTCCLCRFTDNEFHSHSHISTIGVDFKMTIEVDGKVRQ 86
QY 61 IWDTAGQERYQTITKQYRRAGGIFLVYDISERSYQHIMKWSVDVEYAPGVQKILIG 120
DB 87 IWDTAGQERYQTITKQYRRAGGIFLVYDISERSYQHIMKWSVDVEYAPGVQKILIG 146
QY 121 NKADEEQKRVGREGQQQLAKKEYGMDFYETSACTNLNIKESFTRLTVELVLAHRKEBGL 180
DB 147 NKADEEQKRVGREGQQQLAKKEYGMDFYETSACTNLNIKESFTRLTVELVLAHRKEBGL 206
QY 181 RMRASNELALAELEEEGKPEGPANSSKTCWC 212
DB 207 RMRASNELALAELEEEGKPEGPANSSKTCWC 238

RESULT 3
US-09-817-198A-5
; Sequence 5, Application US/09817198A
; Patent No. US20020146758A1
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001188
; CURRENT APPLICATION NUMBER: US/09/817,198A
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Human
US-09-817-198A-5

Query Match      98.8%; Score 1092; DB 10; Length 218;
Best Local Similarity 97.2%; Pred. No. 6.2e-95;
Matches 212; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

QY 1 MAKQYDVLFRLLIGDSVGKTCCLCRFTDNEFHSHSHISTIGVDFKMTIEVDGKVRQ 60
DB 1 MAKQYDVLFRLLIGDSVGKTCCLCRFTDNEFHSHSHISTIGVDFKMTIEVDGKVRQ 60
QY 61 IWDTAGQERYQTITKQYRRAGGIFLVYDISERSYQHIMKWSVDVEYAPGVQKILIG 120
DB 61 IWDTAGQERYQTITKQYRRAGGIFLVYDISERSYQHIMKWSVDVEYAPGVQKILIG 120
QY 121 NKADEEQKRVGREGQQQLAKKEYGMDFYETSACTNLNIKESFTRLTVELVLAHRKEBGL 174
DB 121 NKADEEQKRVGREGQQQLAKKEYGMDFYETSACTNLNIKESFTRLTVELVLAHRKEBGL 174

US-09-817-198A-4
; Sequence 4, Application US/09817198A
; Patent No. US20020146758A1
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001188
; CURRENT APPLICATION NUMBER: US/09/817,198A
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-817-198A-4

Query Match      97.5%; Score 1077; DB 10; Length 212;
Best Local Similarity 97.6%; Pred. No. 1.5e-93;
Matches 207; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDSVGKTCCLCRFTDNEFHSHSHISTIGVDFKMTIEVDGKVRQ 60
DB 1 MAKQYDVLFRLLIGDSVGKTCCLCRFTDNEFHSHSHISTIGVDFKMTIEVDGKVRQ 60
QY 61 IWDTAGQERYQTITKQYRRAGGIFLVYDISERSYQHIMKWSVDVEYAPGVQKILIG 120
DB 61 IWDTAGQERYQTITKQYRRAGGIFLVYDISERSYQHIMKWSVDVEYAPGVQKILIG 120
QY 121 NKADEEQKRVGREGQQQLAKKEYGMDFYETSACTNLNIKESFTRLTVELVLAHRKEBGL 180
DB 121 NKADEEQKRVGREGQQQLAKKEYGMDFYETSACTNLNIKESFTRLTVELVLAHRKEBGL 180
QY 181 RMRASNELALAELEEEGKPEGPANSSKTCWC 212
DB 181 RYCASNELALAELEEEGKTEGPANSSKTCWC 212

RESULT 5
US-09-764-868-1120
; Sequence 1120, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1120
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (139)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (151)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
```

; LOCATION: (161)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (164)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (188)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; US-09-764-868-1120

Query Match 75.3%; Score 832; DB 9; Length 188;
Best Local Similarity 97.6%; Pred. No. 1.4e-70;
Matches 160; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHISTIGVDFKMKTIYVDGKIVRIQ 60
Db 24 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHISTIGVDFKMKTIYVDGKIVRIQ 83
QY 61 INDTAGQERYQTITKQYRRAGIFLYVDISSERSYOHIMKWVSDVDEYAPGVQKILIG 120
Db 84 INDTAGQERYQTITKQYRRAGIFLYVDISSERSYOHIMKWVSDVDEYAPGVQKILIG 143
QY 121 NKADEEQKRVGREGOQOLAKEYGMDFYETSACTNLNIKESFTR 164
Db 144 NKADEEQKRVGREGOQOLAXEYGMDFYETSACTNLNIKESFTR 187

RESULT 6

US-09-925-302-534
; Sequence 534, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 534
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-925-302-534

Query Match 48.1%; Score 532; DB 10; Length 246;
Best Local Similarity 47.6%; Pred. No. 3e-42;
Matches 100; Conservative 49; Mismatches 43; Indels 18; Gaps 3;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHISTIGVDFKMKTIYVDGKIVRIQ 60
Db 44 MAKAYDHLFLLIGDSGVGKTCCLIRPAEDNNYISTIGIDFKIRIVDGGKIKIKQ 103
QY 61 INDTAGQERYQTITKQYRRAGIFLYVDISSERSYOHIMKWVSDVDEYAPGVQKILIG 120
Db 104 VMDTAGQERFKTITAYIRGAMGILVYDITDEKSPENIQNMKSIKENASAGVERLLIG 163
QY 121 NKADEEQKRVGREGOQOLAKEYGMDFYETSACTNLNIKESFTRTLVLAQRKEGL 180
Db 164 NKCDMEARKVQKEQADKLAREHIGIRFFETSAKSMNVDEAFSSLDIL-----LKSG 217
QY 181 RMRASNELALAELEEKGPGPANSKTC 210
Db 218 GRRSGN-----GNKP--PSTDLTCTC 235

RESULT 7

US-09-925-300-1571
; Sequence 1571, Application US/09925300

; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1571
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-925-300-1571

Query Match 48.0%; Score 530.5; DB 10; Length 218;
Best Local Similarity 55.8%; Pred. No. 3.6e-42;
Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;
QY 1 MAKQ-YDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHISTIGVDFKMKTIYVDGKIVRI 59
Db 19 MAKKTYDLLFKLLIGDSGVGKTCVLFREFSDAFTTISTIGIDFKIKVELQGGKKIKL 78
QY 60 QIWDTAGQERYQTITKQYRRAGIFLYVDISSERSYOHIMKWVSDVDEYAPGVQKILI 119
Db 79 QIWDTAGQERFHTTTSYIRGAMGIMLVYDITDKGFSFENISKWLRNIDEHANEDEVKMLL 138
QY 120 GKNKDEEQKRVGREGOQOLAKEYGMDFYETSACTNLNIKESFTRTLVLIQ 171
Db 139 GKNCDMDKRVVPKRGEGQIAREHIGIRFFETSAKANINIEKAFITLAEILR 190

RESULT 8

US-09-794-257-8
; Sequence 8, Application US/09794257
; Patent No. US20020009804A1
; GENERAL INFORMATION:

; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US20020009804A1e1
; FILE REFERENCE: Human G-Proteins
; FILE REFERENCE: 35800/209285
; CURRENT APPLICATION NUMBER: US/09/794,257
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,606
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 207
; TYPE: PRT
; ORGANISM: homo sapiens
; US-09-794-257-8

Query Match 47.8%; Score 528.5; DB 10; Length 207;
Best Local Similarity 54.0%; Pred. No. 5.2e-42;
Matches 95; Conservative 47; Mismatches 33; Indels 1; Gaps 1;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHISTIGVDFKMKTIYVDGKIVRIQ 60
Db 1 MAKTYDVLFRLLIGDSGVGKTCCLFRFSEDAFTTISTIGIDFKIRTIELDGGKIKIQ 60
QY 61 INDTAGQERYQTITKQYRRAGIFLYVDISSERSYOHIMKWVSDVDEYAPGVQKILIG 120
Db 61 INDTAGQERFRTITAYIRGAMGIMLVYDITNEKSFNINRWIRNIEBHASSDVERMILG 120
QY 121 NKADEEQKRVGREGOQOLAKEYGMDFYETSACTNLNIKES-FTRTLVLIQAHRR 175
Db 121 NKCDMDKRVSKERGEKLAIDYGIKRFLETSAKSSANVEAFTFLARDIMTKLNRR 176

```

RESULT 9
US-09-822-860-5
; Sequence 5, Application US/09822860
; Patent No. US20020146795A1
; GENERAL INFORMATION:
; APPLICANT: ZHU, Shaoping et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; FILE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001214
; CURRENT APPLICATION NUMBER: US/09/822,860
; CURRENT FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Discopyge ommata
US-09-822-860-5

Query Match          46.3%; Score 512; DB 10; Length 190;
Best Local Similarity 47.7%; Pred. No. 1.6e-40;
Matches 93; Conservative 52; Mismatches 42; Indels 8; Gaps 1;

QY 6 DVLFRLLIGDSGVGKTCCLCRFTDNEFHHSHISTIGVDFKMKTIYVDGKIVRIQIWDTA 65
DB 1 DYLFKLLIGDSGVGKTCCLFRFSEDAPNTFFISTIGIDFKIRTVELDGKKIKIWDTA 60

QY 66 GOERYQTITKQYRRAGIFLYVDISSERSYQHIMKWSDVDEYAPGQVKILGNRADE 125
DB 61 GOERFRTITAYYRGAMGIMLVYDITNEKSPDNIRNIRNIEEHSADVERMILGNKDM 120

QY 126 EQKRGVGOQOOLAKYGMDFYETSACTNLNIKESFRLVLOAHRKELEGRLMRAS 185
DB 121 NEKQVSKERGEKLAIDYIGIFLETSKASSINVEAFITLARDIMTKLKKH----- 172

QY 186 NELALAELEEEGKP 200
DB 173 NENSLQEAVDKLSP 187

RESULT 10
US-09-834-765-766
; Sequence 766, Application US/09834765
; Patent No. US20020055478A1
; GENERAL INFORMATION:
; APPLICANT: Mary Faris
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Arthur B. Raitano
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT
; TITLE OF INVENTION: AND DETECTION OF CANCER
; FILE REFERENCE: 129.60SUI
; CURRENT APPLICATION NUMBER: US/09/834,765
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/197,647
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 770
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 766
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-765-766

Query Match          45.0%; Score 497; DB 10; Length 162;
Best Local Similarity 55.4%; Pred. No. 3.4e-39;
Matches 87; Conservative 41; Mismatches 29; Indels 0; Gaps 0;

```

```

QY 9 FRLLIGDSGVGKTCCLCRFTDNEFHHSHISTIGVDFKMKTIYVDGKIVRIQIWDTAGOE 68
DB 1 FKLLIGDSGVGKTCVLFRFSEDAPNTFFISTIGIDFKIRTVELDGKKIKIWDTAGOE 60

QY 69 RYQTITKQYRRAGIFLYVDISSERSYQHIMKWSDVDEYAPGQVKILGNRADEQK 128
DB 61 RFTTITAYYRGAMGIMLVYDITNEKSPDNIRNIRNIEEHSADVERMILGNKCDVNDK 120

QY 129 RVQREGOQOOLAKYGMDFYETSACTNLNIKESFRL 165
DB 121 QVSKERGEKLAIDYIGIFLETSKASSINVENAFETL 157

RESULT 11
US-09-967-736-3
; Sequence 3, Application US/09967736
; Patent No. US20020103340A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: RAB PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/967,736
; FILING DATE: 28-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/154,602
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0367 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: LIVRUT04
; CLONE: 2514506
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-967-736-3

Query Match          43.3%; Score 478.5; DB 10; Length 201;
Best Local Similarity 45.0%; Pred. No. 2.5e-37;
Matches 91; Conservative 42; Mismatches 64; Indels 5; Gaps 1;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHHSHISTIGVDFKMKTIYVDGKIVRIQ 60
DB 1 MNPYDVLFRLLIGDSGVGKTCCLLRADFTYTESYSTIGVDFKIRTVELDGKTIKIQ 60

QY 61 IWDTAGQERYQTITKQYRRAGIFLYVDISSERSYQHIMKWSDVDEYAPGQVKILIG 120
DB 61 IWDTAGQERYQTITSSYRGAGIIVYDVTQESYANVKQLQELIDRYASENKNKLVG 120

```


QY 61 IWDTAQERYQITKQYRRAQGIPLVYDISSERSYQHIMKWSDVDEYAPEGVOKILIG 120
Db 81 IWDTAQERFRITTSYYGAHGIIVYDVDTQESFNQVKNQWLQELDRYASENVNKLIVG 140
QY 121 NKADEQKQVGRGQQLAKKEYGMDFYETSACTNLNIKESF 162
Db 141 NKCDLTTKKVVDYTTAKEFADSLGIPFLETSAKNATNVEQSF 182

RESULT 15

US-09-794-257-16
; Sequence 16, Application US/09794257
; Patent No. US2002009804A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 32705, 27423, 27424, 32700, 32712, No. US2002009804A1e1
; FILE REFERENCE: 35800/209285
; CURRENT APPLICATION NUMBER: US/09/794,257
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,606
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pfam accession number PF00071
US-09-794-257-16

Query Match 39.3%; Score 434.5; DB 10; Length 198;
Best Local Similarity 50.9%; Pred. No. 3.3e-33;
Matches 89; Conservative 30; Mismatches 43; Indels 13; Gaps 3;

QY 10 RLLIGDSGVGKTCILCRFTDNEFHSHISTIGVDFKMTIEVDGIVKIRIQIWDTAGQER 69
Db 1 KLVLLIGDSGVGKSSLLIRFTDNKFFVEEYIPTIGVDFTYKTVEVDGKTVKLQIWDTAGQER 60
QY 70 YQITKQYRRAQGIPLVYDISSERSYQHIMKWSDVDEYA--PEGVOKILIGNKAD--- 124
Db 61 FRALRPAYRGAQGLLVYDITSRDSFENVKKWLEELLRHADKDENVPIVLVGNKCDLED 120
QY 125 -----EQKQVGRGQQLAKKEYG-MDFYETSACTNLNIKESFTRITELVLQ 171
Db 121 DEDELTGQKRVVSTEEGELAKELGALPFMETSAKTNTNVEEAFEELAREILK 175

Search completed: June 18, 2003, 15:44:58
Job time : 50 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 23, 2003, 10:45:22 ; Search time 2485 Seconds
(without alignments)
2482.815 Million cell updates/sec

Title: US-09-817-198B-2

Perfect score: 1105

Sequence: 1 MAQYDVLFLLLIGDSVG.....LEEECKPEGPANSSKTCWC 212

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q/cgn2_1/USPTO.spool/US09817198/runat_18062003_144633_27909/app_query.fasta_1.391
-DB=GenEmbl -OPMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCI=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09817198 -CGN_1_1_2496 -runat_18062003_144633_27909 -NCPU=3
-NO_MMALP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_em.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_cm.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pla.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1105	100.0	1054	6	AX399903 Sequence
2	1090	98.6	3139	10	BC027769 Mus muscu
3	1077	97.5	945	10	M83679 Sprague-Daw
4	798	72.2	2560	10	BC013790 Mus muscu
5	546.5	49.5	1613	3	AB006189 Drosophi
6	546.5	49.5	2349	3	AY060425 Drosophi
7	545	49.3	624	9	AF498943 Homo sapi
8	545	49.3	660	9	AF498943 Homo sapi
9	545	49.3	760	4	AF498943 Homo sapi
10	545	49.3	2048	9	BC002977 Homo sapi
11	545	49.3	2818	9	BC002977 Homo sapi
12	541	49.0	1337	10	BC019990 Mus muscu
13	539.5	48.8	765	5	DXGORA2 Discoppye o
14	534.5	48.4	1211	9	BC000799 Homo sapi
15	534.5	48.4	1238	9	HSRAB13 H.sapiens m
16	532	48.1	612	9	AF498948 Homo sapi
17	531.5	48.1	840	5	DYGORAI Discoppye o
18	530.5	48.0	603	9	AF297660 Homo sapi
19	530.5	48.0	603	9	AF498945 Homo sapi
20	530.5	48.0	888	4	CFRAB10 Canine rabl
21	530.5	48.0	897	10	AF035646 Mus muscu
22	530.5	48.0	1029	9	BC000896 Homo sapi
23	530.5	48.0	3164	9	AF106681 Homo sapi
24	530.5	48.0	3533	9	AK023223 Homo sapi
25	529.5	47.9	954	8	ATHARA3 Arabidopsi
26	529.5	47.9	969	8	AY042795 Arabidopsi
27	529.5	47.9	1143	8	AY035132 Arabidops
28	528.5	47.8	624	6	AX236078 Sequence
29	528.5	47.8	740	10	U53475 Rattus norv
30	528.5	47.8	759	10	AF525280 Rattus nor
31	528.5	47.8	1128	9	BC020654 Homo sapi
32	528.5	47.8	1161	6	AX236076 Sequence
33	528.5	47.8	1265	9	AB038995 Homo sapi
34	528.5	47.8	2048	3	AY069671 Drosophi
35	528.5	47.8	2210	3	D84347 Drosophila
36	528.5	47.8	2497	6	AX285074 Sequence
37	528.5	47.8	2497	6	AX285080 Sequence
38	528.5	47.8	2497	6	AX285089 Sequence
39	528.5	47.8	2497	6	AX285091 Sequence
40	528.5	47.8	3077	9	AK001111 Homo sapi
41	528.5	47.8	3110	9	AL833365 Homo sapi
42	526.5	47.6	1113	8	LJRA88E L.japonicus
43	526.5	47.6	3521	9	AF136650 Homo sapi
44	526	47.6	638	10	S53270 MEL-RAS-rel
45	526	47.6	1118	10	BC027214 Mus muscu

ALIGNMENTS

RESULT 1

AX399903 LOCUS AX399903 1054 bp DNA linear PAT 06-JUN-2002
 DEFINITION Sequence 74 from Patent WO2018424.
 ACCESSION AX399903
 VERSION AX399903.1 GI:21336157
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 REFERENCE Tang,Y.T., Asundi,V., Zhou,P., Xue,A.J., Ren,F., Zhang,J.,
 AUTHORS Wang,J.R., Zhao,Q.A., Wang,D., Liu,C., Drmanac,R.T. and Wehrman,T.
 TITLE Nucleic acids and polypeptides
 JOURNAL Patent: WO 0218424-A 74 07-MAR-2002;
 HYSEQ, INC. (US)
 FEATURES
 source Location/Qualifiers
 1..1054
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 81..719
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD33347.1"
 /db_xref="GI:21336158"
 /translation="WAKQYDVLFRLLIGDSGVGKTCLLCRFTDNFHHSHISTIGVD
 FMKTIIEVGIKVRQIWMQTAQERYQTITKQYRAQGIFLVYDITSSRSYOHIMKW
 VSDDEYAPEGVKILIGNKADEOKRQVREGQOLAKEYGMDFYETACTNLNIKE
 SFTRTELVLQAHKELEGRLRNSELALAELEEEGPEGPANSSKTCWC"
 BASE COUNT 240 a 314 c 304 g 195 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,15e-100 Length: 1054
 Score: 1105.00 Matches: 212
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-817-198B-2 (1-212) x AX399903 (1-1054)
 QY 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyValGly 20
 |||||
 Db 81 ATGGCGAAGCAGTACGATGCTGTTCGCGCTGCTGCTGATCGGGACTCCGGGTGGC 140
 |||||
 QY 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisIleSerThr 40
 |||||
 Db 141 AAGACCTGCGTGTGTGCGCGCTTCACCGAACAGAGTTCACCTCCGCGACATCTCCAC 200
 |||||
 QY 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60
 |||||
 Db 201 ATCGGTGTGACTTTAAGATGAAGACCATAGAGGTAGACGGCATCAAGTGGGATACAG 260
 |||||
 QY 61 IleTrpAspThrAlaGlyGlnArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
 |||||
 Db 261 ATCTGGGACACTCGACGGCAGGAGATACACGACCATCACAAGCAGTACTATCGCGG 320
 |||||
 QY 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
 |||||
 Db 321 GCCCGGGGATATTTTGGTCTATGATATAGCAGGCGGCTCTTACCAGCACATCATG 380
 |||||
 QY 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
 |||||
 Db 381 AAGTGGGTCAGTGACGTGATGAGTACGACCGACCGAAGGGCTCCAGAAGATCTTATTGG 440
 |||||
 QY 121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAla 140
 |||||
 Db 441 AATAAGCCTCATGAGGAGCAGAAACCGCAGGTGGGAAGAGAGCAAGGGCAGCAGCTGGCG 500
 |||||
 QY 141 LysGluTyrGlyMetaspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
 |||||
 Db 501 AAGGAGTATGGGATGACTTCTATGAACAGTGGCTGCACCACTCAACACTTAAGAG 560
 |||||

Qy	161	SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu	180
Db	561	TCATTACACGGCTGTGACAGAGCTGGTGCTGCAGGCCCACATAGGAGGAGCTGGAGGCCTC	620
Qy	181	ArgMetArgAlaSerIasnGluLeuAlaLeuAlaGluLeuGluGluGlyLysPro	200
Db	621	CGATGCGTCCACGACATAGTTGGCACTGGCAGAGCTGGAGGAGGAGGAGGCAAAACC	680
Qy	201	GluGlyProAlaAsnSerSerLysThrCysTrpCys	212
Db	681	GAGGGCCCCGGAACCTCTTCGAAACCTGCTGGTGC	716
RESULT 2			
LOCUS	BC027769	3139 bp	mRNA linear ROD 07-AUG-2002
DEFINITION	Mus musculus, clone MGC:38375 IMAGE:5345297, mRNA, complete cds.		
ACCESSION	BC027769		
VERSION	BC027769.1 GI:20380721		
KEYWORDS	MGC.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus; 1 (bases 1 to 3139)		
AUTHORS	Strausberg,R.		
TITLE	Direct Submission		
JOURNAL	Submitted (08-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk Email: cgabps-r@mail.nih.gov Tissue Procurement: Jeffrey Green M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc.mgc@nih.gov Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., BenJamin, B., Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Lari, P., Legaspi, R., Maduro, O.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgou, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.		

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAK Plate: 55 Row: 4 Column: 5
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES	SOURCE
-----------------	---------------

```

location/Qualifiers
1..3139
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:38375 IMAGE:5345297"
/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
ductal carcinoma, 5 month old virgin mouse."
/clone_lib="NCI-CCAP_Mam6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
208...846
/codon_start=1
/product="Unknown (protein for MGC:38375)"
/protein_id="AAH27769.1"
/db_xref="GI:20380722"

```

CDS

161	Qy	SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu	180
700	Db	TCCTTCACCTGCTGTGACCGAGCTGGTCTGCAGGCTCACAGAAAGAGCTGGATGGTCTC	759
181	Qy	ArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGluGluGlyLysPro	200
760	Db	CGAACATGTGCCAGCAATGAGTCGCACCTGSCCGAGCTGGAGGAAGACGAAGGCAAAACT	819
201	Qy	GluGlyProAlaAsnSerSerLysThrCysTrpCys	212
820	Db	GAAGGGCCCAACAACTCTTCAAGACACCTGTCTGGTGC	855

RESULT 4	BC013790	2560 bp	linear	ROD 07-AUG-2002
LOCUS	Mus musculus	clone MGC:6897	IMAGE:2655151	mRNA, complete cds.
DEFINITION	BC013790			
ACCESSION	BC013790.1	GI:15489933		
KEYWORDS	MGC			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	1 (bases 1 to 2560)			
AUTHORS	Strausberg, R.			
TITLE	Direct Submission			
JOURNAL	Submitted (04-SEP-2001) National Institutes of Health, Mammalian			
	Gene Collection (MGC), Cancer Genomics Office, National Cancer			
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,			
	USA			
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov			

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 5 Row: 1 Column: 20
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Genomescan gene prediction.

[illegible]

**BASE COUNT
ORIGIN**

REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	1 (bases 1 to 660)
TITLE	Zahraoui,A.
JOURNAL	Direct Submission
-	Submitted (26-NOV-1990) A. Zahraoui, INSERM-U 248, 10 AVENUE DE VERDUN, 750-10 PARIS, FRANCE
REFERENCE	2 (bases 1 to 660)
AUTHORS	Zahraoui,A., Joberty,G., Arpin,M., Fontaine,J.J., Helllo,R., Tavitian,A. and Louvard,D.
TITLE	A small rab GTPase is distributed in cytoplasmic vesicles in non polarized cells but colocalizes with the tight junction marker ZO-1 in polarized epithelial cells
JOURNAL	J. Cell Biol. 124 (1-2), 101-115 (1994)
MEDLINE	94124602
PUBMED	8294494
FEATURES	
SOURCE	Location/Qualifiers
	1..660
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
.gene	/clone_lib="Human pheochromocytoma cDNA library"
	11..634
	/gene="rab8"
CDS	11..634
	/gene="rab8"
	/codon_start=1
	/product="rab8 small GTP binding protein"
	/protein_id="CAA40065.1"
	/db_xref="GI:452318"
	/db_xref="SWISS-PROT:P24407"
	/translation="MAKTYDIYLFKLIIIGDSVGKTGVLPFRSEDAFNFTFTTIGID FKIRIEDGKRKIQLWDTAGQRFRTITTAAYRGAMGIMLYVIDTNKESFDNINRW INRIEHSADYKEMILGNCKVDNRQVKRGEKALDYGIKFMTSAKANINVEN AFETLARDIKAKMDKKLEGNPQGSNOGVKITPDQOKRSFFRCVL"
BASE COUNT	197 a 159 c 172 g 132 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3,64e-45 Length: 660
Score:	545.00 Matches: 99
Percent Similarity:	77.01% Conservative: 45
Best Local Similarity:	52.94% Mismatches: 43
Query Match:	49.32% Indels: 0
DB:	Gaps: 0
US-09-817-198B-2 (1-212) x HSMRAB8 (1-660)	
QY	1 MetAlaLysGlnTyAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyValGly 20
Db	11 ATGGCGAAGACCTTAGCATTACCTGTTCATGTTCAAGCTGCTGCTGATCGGGACTCGGGGTGGG 70
QY	21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisIleSerThr 40
Db	71 AGACCTGTGCTGTTCCGCTTCGAGGAGGCCCTTCAACCHCACCTTTATCTCCACC 130
QY	41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyLysValArgIleGln 60
Db	131 ATAGGAATGACTTTAAAAATTAGGACCATAGAGCTCGATGGCAAGAATAAACTGCAG 190
QY	61 IleTrpAspThrAlaGlyGlnGluArgTyrrGlnThrIleThrLysGlnTyrrArgArg 80
Db	191 ATATGGGACACAGCCCGGTGAGAACGGTTTCGACGATCACACAGCCCTACTACAGGGGT 250
QY	81 AlaGlnGlyIlePheLeuValTyrrAspIleSerSerGluArgSerTyrrGlnHisIleMet 100
Db	251 GCATGGGGCATCATGCTGTGTACGACATCCACGAGAAAGTCTTCGACAACATCCCG 310
QY	101 LysTrpValSerAspValAspGlnTyrrAlaProGluGlyValGlnLysIleLeuIleGly 120
Db	311 AACCTGGATTCCGAACATCTCAGGAGCAGCGCTCTGTCAGACGTCGAAAAGATGATCTCGG 370
QY	121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAla 140
Db	371 ACAACAGTGTCGTGAATCACAGACAGCAAGTTTCCAAGGACGGGCGAAGAACTGGCC 430

QY	141	LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu	160
Db	431	CTCGACTATGGAATCAAGTTTCATGGACGACGCGGCAACATCAATGTGAAAAAT	490
QY	161	SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu	180
Db	491	GCATTTTCACCTCTGCCAGAGATATCAAGCAAAATGGACAAAAATTTGAAGGCAAC	550
QY	181	ArgMetArgAlaSerAsnGlu	187
Db	551	AGCCCCCAGGGGAGCAACACG	571
RESULT 9			
CFRAB8			
LOCUS		Canine rab8 mRNA for ras-related GTP-binding protein.	
DEFINITION		Canine rab8 mRNA for ras-related GTP-binding protein.	
ACCESSION		X56385	
VERSION		X56385.1 GI:920	
KEYWORDS		GTP-binding protein; protein transport; ras gene; ras gene family.	
SOURCE		Canis familiaris.	
ORGANISM		Canis familiaris.	
REFERENCE			
AUTHORS		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
TITLE		Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.	
JOURNAL		1 (bases 1 to 760)	
MEDLINE		Chavrier,P., Vingron,M., Sander,C., Simons,K. and Zerial,M.	
PUBMED		Molecular cloning of YP11/SEC4-related cDNAs from an epithelial	
REFERENCE		cell line	
AUTHORS		Mol. Cell. Biol. 10 (12), 6578-6585 (1990)	
TITLE		91061765	
JOURNAL		2123294	
MEDLINE		2 (bases 1 to 760)	
PUBMED		Zerial,M.	
REFERENCE		Direct Submission	
AUTHORS		Submitted (17-AUG-1990) Zerial M., EMBL, Meyerhofstrasse 1, 6900	
TITLE		Heidelberg, FRG	
JOURNAL			
FEATURES			
source		Location/Qualifiers	
		1..760	
		/organism="Canis familiaris"	
		/strain="cocker spaniel"	
		/db_xref="taxon:9615"	
		/clone="2"	
		/cell_line="MDCK"	
gene		10..633	
		/gene="rab8"	
CDS		10..633	
		/gene="rab8"	
		/function="GTP binding protein"	
		/codon_start=1	
		/protein_id="CAB56776.1"	
		/db_xref="GI:6006436"	
		/translation="MAKTYDYLFKLLIGDSGVGKTCVLFRESEDAFNSTFTSTIGID	
		FKRTIESADYDKRIKIQIWDACQGRFTTITAYRGAMGIMLYDITNEKSPDNIRNW	
		IRNIIEHSADEKVMILGNQVNDKROVSKERGEKALDYGIKEMETSAKANINVEN	
		AFETLARDIKARMDKKLEGNSPQGSNOGVKTIPTDQKRSPFFRCVLL"	
BASE COUNT		215 a 190 c 196 g 159 t	
ORIGIN			
Alignment Scores:			
Pred. No.:		4.32e-45	Length: 760
Score:		545.00	Matches: 99
Percent Similarity:		77.01%	Conservative: 45
Best Local Similarity:		52.94%	Mismatches: 43
Query Match:		49.32%	Indels: 0
DB:		4	Gaps: 0
US-09-817-198B-2 (1-212) x CFRAB8 (1-760)			
QY	1	MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuGlyAspSerglyValGly	20
Db	10	ATGGCGAAGACCTCATTTACCTGTTCACAGCTGCTCTCATCGGGACTCGGGGTGGG	69
QY	21	LysThrCysLeuLeuCvsArgPheThrAsoAsnGluPheHisSerSerHisIleSerThr	40

```

Db 70 AAGACCTGTGCTGCTCCGAGGAGCGCTTCAACTCAACTTTCATCTCCACT 129
QY 41 IleglyValAspPheLysMetLysThrIleGluValAspGlyLeuValArgGln 60
Db 130 ATAGGAATTGACTTTAAATAGGACCATAGAGCTGATGCGAAGAGATTAACTACAG 189
QY 61 IletPaspThralaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArg 80
Db 190 ATATGGACACAGCTGCTCAAGAACGTTTTCGAGCATCATCAACAGCTTATACAGGCG 249
QY 81 AlaGlnGlyPheLeuValTyrAspPheSerSerGluArgSerTyrGlnHisIleMet 100
Db 250 GCAATGGGATCATGCTGCTATGACATCAACAGAGAAATCTTTCACATATCCGG 309
QY 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
Db 310 AACTGGATTCGGAACATTGAGGAGCATGCTTCCGACATGTCGAAAGATGATACCTCGA 369
QY 121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAla 140
Db 370 AACAGTGTGATGTGACGACAAAGACAAAGTTTCCAGGAACGGGGAGAAAGCTGGCC 429
QY 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
Db 430 CTGGACTATGGATCAAGTTATGAGACCATGCGAAGCGCAACATCAATGTGGAGAAC 489
QY 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluLeu 180
Db 490 GCATTTTTCATCTCCGACAGACATCAAGCAAGATGCGACAAAGAAATTTGGAAGGCAAC 549
QY 181 ArgMetArgAlaSerAsnGlu 187
Db 550 AGTCCCAAGGAGCAACCAAG 570

RESULT 10
BC002977 2048 bp mRNA linear PRI 12-JUL-2001
LOCUS Homo sapiens, mel transforming oncogene (derived from cell line
DEFINITION NK14)- RAB8 homolog, clone MGC:2196 IMAGE:3547214, mRNA, complete
cds.
ACCESSION BC002977 GI:12804236
VERSION MGC.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Cranial; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2048)
Strausberg, R.
Direct Submission
Submitted (06-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DCMP/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
Info@bgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeary, Steven
Ness, Pawan Pandon, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schlein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 8 Row: d Column: 22
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 452317.

FEATURES
 Location/Qualifiers
 1..2048
 /organism="Homo sapiens"
 /db_xref="LocusID:4218"
 /db_xref="taxon:9606"
 /clone="MGC:2196 IMAGE:3547214"
 /tissue_type="Skin, melanotic melanoma."
 /clone_lib="NIH_MGC_20"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 77..700
 /codon_start=1
 /product="mel transforming oncogene (derived from cell line NK14)- RAB8 homolog"
 /protein_id="AAH02977.1"
 /db_xref="GI:12804237"
 /translation="MAKTYDYLFKLLILIGDSGVGKTVLFRSEDAFNSTFTSTIGIDFKRTIELDGKRIKLIQINDTAGQERFRTTAYYRGAMGIMLVYDITNEKSPDNLRNWLENIHEHADVEKMLGNKCDYNDKQVSKERKGLALDYGIKIPWESAKANINVENAFPTLARDIKAKDKLDENSPQNGQVKIITPDQOKRSSFFRCVLL"

BASE COUNT 532 a 556 c 493 g 467 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1,45e-44 Length: 2048
 Score: 545.00 Matches: 99
 Percent Similarity: 77.01% Conservative: 45
 Best Local Similarity: 52.94% Mismatches: 43
 Query Match: 49.32% Indels: 0
 DB: 9 Gaps: 0

US-09-817-198b-2 (1-212) x BC002977 (1-2048)

QY 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyValGly 20
 Db 77 ATGGGAGACCTACGATTACCTTCAAGCTGCTGATCGGGGACTCGGGGGTGGG 136
 QY 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisIleSerThr 40
 Db 137 AAGACCTGTGCTGCTCCGAGGAGCGCTTCACTCCACTTTTATCTCCACC 196
 QY 41 IleglyValAspPheLysMetLysThrIleGluValAspGlyLeuValArgGln 60
 Db 197 ATAGGAATTGACTTTAAATAGGACCATAGAGCTGATGCGAAGAGATTAACTGCGAG 256
 QY 61 IletPaspThralaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArg 80
 Db 257 ATATGGACACAGCGGTCAGAACGTTTTCGAGCATCATCAACGCGCTACTACAGGGT 316
 QY 81 AlaGlnGlyPheLeuValTyrAspPheSerSerGluArgSerTyrGlnHisIleMet 100
 Db 317 GCAATGGGATCATGCTGCTTACGACATCAACAGAGAGTCTTCCACACATATCCGG 376
 QY 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
 Db 377 AACTGGATTCGGAACATTGAGGAGCATGCTTCCGAGCATCTGCGAGAGATGATACCTCGG 436
 QY 121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAla 140
 Db 437 AACAGTGTGATGTAATGACAAGACAAAGTTTCCAGGAACGGGGAGAAAGCTGGCC 496
 QY 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
 Db 497 CTCGACTATGGATCAAGTTATGAGACCATGCGAAGCGCAACATCAATGTGGAAAT 556
 QY 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluLeu 180
 Db 557 GCATTTTTCATCTCCGACAGATATCAAGCAAGAAATGGACAAAATTTGGAAGGCAAC 616

Qy 181 ArgMetArgAlaSerAsnGlu 187
Db 617 AGCCCCCAGGGGAGCAACAC 637

RESULT 11
AK025165.
LOCUS Homo sapiens cDNA: FLJ21512 fis, clone COL05769. PRI 29-SEP-2000
DEFINITION AK025165
VERSION AK025165.1 GI:10437627
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens colon cDNA to mRNA, clone_lib.COL clone:COL05769.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2818)
AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
DIRECT SUBMISSION
TITLE Submitted
JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).

FEATURES
source Location/Qualifiers
1 .2818
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COL05769"
/tissue_type="colon"
/clone_lib="COL"
/note="cloning vector pME18SFL3"
BASE COUNT 701 a 800 c 708 g 609 t
ORIGIN

Alignment Scores:
Pred. No.: 2.15e-44 Length: 2818
Score: 545.00 Matches: 99
Percent Similarity: 77.01% Conservative: 45
Best Local Similarity: 52.94% Mismatches: 43
Query Match: 49.32% Indels: 0
DB: 9 Gaps: 0

US-09-817-198B-2 (1-212) x AK025165 (1-2818)

Qy 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuGlyAspSerGlyValGly 20
Db 246 ATGGCGAAGACCTACGATACCTGTTCAAGCTGCTGCTGATCGGGGACTCGGGGTGGG 305

Qy 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSerThr 40
Db 306 AAGACCTGTGCTGCTTCGGCTTCCGAGACGCCCTCACTCCACTTTATCTCCACC 365

Qy 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60
Db 366 ATAGGAATTGACTTTAAATTTAGGACCATAGCTCGATGTCGAAGAGAAATTAACCTGCAG 425

Qy 61 IleTrpAspThrAlaGlnGluArgTyrGlnThrIleThrLysGlnTyrArgArg 80
Db 1

Db 426 ATATGGGCACACGCCGTCAGGAACGGTTTCGGAGCATCAACAGGCCCTACTACAGGGGT 485

Qy 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
Db 486 GCAATGGGCATCATGCTGCTACGACATCAACCAAGAGAGAGTCTCTTCGACCAACATCGG 545

Qy 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
Db 546 AACGTGGATTCCCAACATTTAGGAGCAGCCCTCTCGACAGCTCGAAAAGATGATATCGGG 605

Qy 121 AsnLysAlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGlnLysAla 140
Db 606 AACAGGTGTGATGTGAATGACAGACAGCAAGTTTCCAGAGGACGGGAGAAAGCTGCC 665

Qy 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
Db 666 CTCGACTATGGAATCAAGTTTCATGAGCAGCAGCGAAGGCAACATCAATGTGGAATAAT 725

Qy 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu 180
Db 726 GCATTTTTCATCTCTCGCAGAGATATCAAGCAAAAAATGGACAAAAATTCGAAGGCAAC 785

Qy 181 ArgMetArgAlaSerAsnGlu 187
Db 786 AGCCCCCAGGGGAGCAACAC 806

RESULT 12
LOCUS BC019990
DEFINITION Mus musculus, Similar to mel transforming oncogene (derived from
cell line NK14) - RAB8 homolog, clone MGC:28462 IMAGE:4161110, mRNA,
complete cds.
ACCESSION BC019990
VERSION BC019990.1 GI:18043408
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1337)
TITLE Direct Submission
JOURNAL Strausberg,R.
Submitted (19-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunnaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAP Plate: 37 Row: c Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.
Location/Qualifiers
1 .1337
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:28462 IMAGE:4161110"

```

/issue_type="Liver, normal. 5 month old male mouse."
/clone_lib="NCI_CGAP_f19"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. 624
/codon_start=1
/product="Similar to mel transforming oncogene (derived
from cell line NK14)- RAB8 homolog"
/protein_id="AAH19990.1"
/db_xref="GI:18043409"
/translation="WAKTVDFLKLLIGDSGVGTCVLFRESEDAFNSTFISICID
FIRIELDGKRIKLIWDTAGQDERPITTTAYRGAGMILVYDITNEKSPDNIRN
INIEHSEADVEKILMGNDKDRKRGKEKALDYGKTFMETSAKANINVEN
AFPTFLADTKAKMDKLEGNPSQGGSHVKYITVEQQKRTSFRCSLI"
367 a 338 c 341 g 291 t
BASE COUNT

```

Alignment Scores:		
Pred. No.:	2.15e-44	Length:
Score:	541.00	Matches:
Percent Similarity:	75.13%	Conservative:
Best Local Similarity:	50.25%	Mismatches:
Query Match:	48.96%	Indels:
DB:	10	Gaps:
US-09-817-198B-2 (1-212) x BC019990 (1-1337)		

Db	440	AACTGGATGAAAACGATCAGGAGAATGCCTCGGCTGGGTGAGCGCCTCTGCTGGGG	499
QY	121	AsnLysAlaAspGluGluGlnLysArgGlnValGlyArgGlnGlnGlyGlnGlnLeuAla	140
Db	500	AACAAATGTGACATGGAGGCCAAGAGGAAAGTGCAGNAGGAGCAGGCCCATAGTTGGCT	559
QY	141	LysGluIyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu	160
Db	560	CGAGAGCATGGAATCCGATTTTTCGAAACATAGTCTAAATCCAGTATGAATCGGTAG	619
QY	161	SerPheThrArgLeuThrGluLeuValLeu-----	170
Db	620	GCTTTTAGTTCCTGCGCCGGGACATCTGCTCAAGTCAGGAGCGCGAGATCAGGAAC	679
QY	171	-----GlnAlaHisArgLysGlu	176
Db	680	GGCAACAAGCCTCCAGTACTGTACCTGAAACCTGTGTACAAGAAGAA-CACCAACAAGTG	738
QY	177	LeuGluGlyLeuArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGlu	196
Db	739	CTCCCTGGCTGAGGACCCCTTTCITGCTCCCAACCCCGGAAGCTCAACTGAGGGAGAC	798
QY	197	GluGlyLysProGluGlyProAlaAsnSerSerLysThrCysTrp	211
Db	799	AACGGCAGAGGAGTGTAGCAGGGGAGAGAAATAGCAGAGGGGCTTGG	843

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 23, 2003, 03:05:27 ; Search time 227 Seconds
(without alignments)
2103.188 Million cell updates/sec

Title: us-09-817-198b-2

Perfect score: 1105

Sequence: 1 MAKQDVLFRLLLIGDVG.....LEERKPEGPANSSKTCWC 212

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO/spool/US09817198/runat_18062003_144632_27901/app_query.fasta_1.391
-DB=N_Geneseq_101002 -QPMF=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09817198 -ECGN_1_1_263 -runat_18062003_144632_27901 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_101002: *
1: /SDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT: *
2: /SDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT: *
3: /SDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT: *
4: /SDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT: *
5: /SDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT: *
6: /SDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT: *
7: /SDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT: *
8: /SDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT: *
9: /SDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT: *
10: /SDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT: *
11: /SDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT: *
12: /SDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT: *
13: /SDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT: *
14: /SDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT: *
15: /SDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT: *
16: /SDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT: *
17: /SDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT: *
18: /SDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT: *
19: /SDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT: *
20: /SDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT: *
21: /SDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT: *
22: /SDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT: *
23: /SDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT: *
24: /SDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	1105	100.0	2021	22	AAS27053
2	1092	98.8	1666	21	AAS27053
3	832	75.3	566	22	AA75813
4	546.5	49.5	1540	23	ABL27707
5	545	49.3	1986	22	ABA09160
6	540	48.9	2411	24	ABO54410
7	534.5	48.4	1274	21	AAF18072
8	530.5	48.0	716	21	AAA40104
9	530.5	48.0	861	21	AAA40108
10	530.5	48.0	888	21	AAA40109
11	530.5	48.0	956	21	AAA96887
12	530.5	48.0	1537	21	AAF16196
13	530.5	48.0	3533	22	AAH17889
14	529.5	47.9	674	23	AAS71453
15	528.5	47.8	1161	22	AAH75182
16	528.5	47.8	2247	23	ABL29661
17	528.5	47.8	2497	22	AAS60878
18	528.5	47.8	2497	22	AAS60884
19	528.5	47.8	2497	22	AAS60893
20	528.5	47.8	2497	22	AAS60895
21	528.5	47.8	2497	23	ABV25781
22	528.5	47.8	2497	23	ABV30037
23	528.5	47.8	3077	22	AAH13912
24	519.5	47.0	866	22	AAH04301
25	519	47.0	1023	21	AAC51491
26	519	47.0	1025	21	AAC34347
27	513.5	46.5	1129	21	AAC44482
28	510	46.2	911	21	AAC38429
29	507	45.9	1101	21	AAC34080
30	504	45.6	1203	21	AAC34385
31	501	45.3	1705	21	AAC49208
32	494	44.7	1153	21	AAC32781
33	478.5	43.3	925	20	AAZ27232
34	478.5	43.3	939	22	AAZ21661
35	478.5	43.3	1898	22	AAS44924
36	478.5	43.3	1944	22	AAI93456
37	477.5	43.2	730	21	AAC36696
38	477.5	43.2	777	21	AAC42684
39	477.5	43.2	881	24	ABN98761
40	472.5	42.8	607	22	AAH87926
41	472.5	42.8	1202	21	AAF14053
42	468	42.4	959	21	AAC33987
43	467	42.3	723	24	ABR83378
44	467	42.3	2528	21	AAC60009
45	466	42.2	932	21	AAC35200

ALIGNMENTS

RESULT 1

AAS27053

ID AAS27053 standard; cdna; 2021 BP.

XX AAS27053;

XX 07-NOV-2001 (first entry)

XX cdna encoding novel signal transduction pathway protein, Seq ID 88.

XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
XX antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
XX immune system disorder; rheumatoid arthritis; inflammatory condition;
XX organ transplant rejection; infection; hepatitis C; blood disorder;
XX sickle cell anemia; hyperproliferative disorder; Gaucher's disease;
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
XX cardiovascular; respiratory; wound healing; endocrine; Addison's disease;

KW reproductive system; gastrointestinal; liver disorder; AIDS; ss;
KW acquired immune deficiency syndrome.

XX Homo sapiens.

OS WO200154733-A1.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01312.

XX PR 31-JAN-2000; 2000US-0179065.

XX PR 04-FEB-2000; 2000US-0180628.

XX PR 24-FEB-2000; 2000US-0184664.

XX PR 02-MAR-2000; 2000US-0186350.

XX PR 16-MAR-2000; 2000US-0189874.

XX PR 17-MAR-2000; 2000US-0190076.

XX PR 18-APR-2000; 2000US-0198123.

XX PR 19-MAY-2000; 2000US-0205515.

XX PR 07-JUN-2000; 2000US-0209467.

XX PR 28-JUN-2000; 2000US-0214886.

XX PR 30-JUN-2000; 2000US-0215135.

XX PR 07-JUL-2000; 2000US-0216647.

XX PR 07-JUL-2000; 2000US-0216880.

XX PR 11-JUL-2000; 2000US-0217487.

XX PR 11-JUL-2000; 2000US-0217496.

XX PR 14-JUL-2000; 2000US-0218290.

XX PR 26-JUL-2000; 2000US-0220963.

XX PR 26-JUL-2000; 2000US-0220964.

XX PR 14-AUG-2000; 2000US-0224518.

XX PR 14-AUG-2000; 2000US-0224519.

XX PR 14-AUG-2000; 2000US-0225213.

XX PR 14-AUG-2000; 2000US-0225214.

XX PR 14-AUG-2000; 2000US-0225266.

XX PR 14-AUG-2000; 2000US-0225267.

XX PR 14-AUG-2000; 2000US-0225268.

XX PR 14-AUG-2000; 2000US-0225270.

XX PR 14-AUG-2000; 2000US-0225447.

XX PR 14-AUG-2000; 2000US-0225757.

XX PR 14-AUG-2000; 2000US-0225758.

XX PR 14-AUG-2000; 2000US-0225759.

XX PR 18-AUG-2000; 2000US-0226279.

XX PR 22-AUG-2000; 2000US-0226681.

XX PR 22-AUG-2000; 2000US-0226688.

XX PR 22-AUG-2000; 2000US-0227182.

XX PR 23-AUG-2000; 2000US-0227009.

XX PR 30-AUG-2000; 2000US-0228924.

XX PR 01-SEP-2000; 2000US-0229287.

XX PR 01-SEP-2000; 2000US-0229343.

XX PR 01-SEP-2000; 2000US-0229344.

XX PR 01-SEP-2000; 2000US-0229345.

XX PR 05-SEP-2000; 2000US-0229509.

XX PR 05-SEP-2000; 2000US-0229513.

XX PR 06-SEP-2000; 2000US-0230437.

XX PR 06-SEP-2000; 2000US-0230438.

XX PR 08-SEP-2000; 2000US-0231242.

XX PR 08-SEP-2000; 2000US-0231243.

XX PR 08-SEP-2000; 2000US-0231244.

XX PR 08-SEP-2000; 2000US-0231413.

XX PR 08-SEP-2000; 2000US-0231414.

XX PR 08-SEP-2000; 2000US-0232080.

XX PR 08-SEP-2000; 2000US-0232081.

XX PR 12-SEP-2000; 2000US-0231968.

XX PR 14-SEP-2000; 2000US-0232397.

XX PR 14-SEP-2000; 2000US-0232398.

XX PR 14-SEP-2000; 2000US-0232399.

XX PR 14-SEP-2000; 2000US-0232400.

XX PR 14-SEP-2000; 2000US-0232401.

XX PR 14-SEP-2000; 2000US-0233063.

XX PR 14-SEP-2000; 2000US-0233064.

XX PR 14-SEP-2000; 2000US-0233065.

XX PR 21-SEP-2000; 2000US-0234223.

XX PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249266.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.

PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231245.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246533.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.

PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249285.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX P-PSDB: 2001-465460/50.
 XX P-PSDB: AAU17555.

Novel polypeptides useful for diagnosing, treating, preventing and/or
 prognosing disorders related to the proteins, including cancers, immune
 disorders and neuronal disorders

Claim 1; SEQ ID No 507; 880pp; English.

The invention relates to novel isolated polypeptides (I), and
 polynucleotides (II). (I), (II) and the antibody to (I) are useful for
 diagnosing, preventing and treating diseases including immune system
 disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
 disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
 transplant rejections and graft versus host disease, infectious diseases
 (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
 other blood-related disorders (sickle cell anaemia), myeloproliferative
 disorders, primary haematopoietic disorders, hyperproliferative
 disorders (e.g. Gaucher's disease and cancer), neurodegenerative
 disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
 abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
 disorders (e.g. glomerulonephritis), cardiovascular disorders
 (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
 wound healing, epithelial cell proliferation, endocrine disorders (e.g.
 Addison's disease), reproductive system disorders, gastrointestinal
 disorder (inflammatory disorders), liver disorders (cirrhosis),
 as stimulators of B-cell responsiveness to pathogens, activators of
 T-cells, to induce higher affinity antibodies, and as a means to induce
 tumour proliferation in pathologies e.g. acquired immune deficiency
 syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction
 pathway protein coding sequences and PCR primers of the invention.

Alignment Scores:

Pred. No.:	4.6e-87	Length:	566
Score:	832.00	Matches:	160
Percent Similarity:	97.56%	Conservative:	0
Best Local Similarity:	97.56%	Mismatches:	4
Query Match:	75.29%	Indels:	0
DB:	22	Gaps:	0

US-09-817-198B-2 (1-212) x AAS27472 (1-566)

QY	1	MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuGlyAspSerGlyValGly	20
Db	71	ATGCCGAAGCAGTACGATGTCTGCTCGGCTGCTGTGCTGCTCGGGACTCCGGGTGGC	130

XX	Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW	haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW	inhibin; chemotaxis; chemokines; thrombolysis; oncogenesis;
KW	proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW	myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW	chronic inflammatory condition; proliferative retinopathy;
KW	atherosclerosis; coronary heart disease; arterial ischaemia;
KW	bone disorder; osteoporosis; vascular growth disorder;
KW	tissue regeneration; wound healing; infection; immune disorder;
KW	cell culture; drug screening; gene therapy; antiinflammatory;
KW	antiallergic; antiarthritic; haemostatic; antiarteriosclerotic;
KW	cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW	antifungal; vulvarny; antitumor; ss.

Homo sapiens.

WO200157188-A2.

09-AUG-2001.

05-FEB-2001: 2001WO-US03800.

03-FEB-2000: 2000US-0496914.

03-FEB-2000; 2000US-0496914;
27-APR-2000; 2000US-0560875;

(HYSE-) HYSEO INC.

Tang YT, Liu C, Drmanac RT:

WPI: 2001-457740/49.

P-PSDB; ABB11916.

Human proteins and DNA encoding sequences useful for preventing, treating or ameliorating a medical condition in a mammalian subject

Claim 1; Page 805; 1963pp: English.

Sequences AB010901-AB012330 represent 1350 novel human polypeptides, and sequences ABA08225-ABA09574 represent nucleic acids encoding them. The invention also relates to vectors and recombinant host cells comprising a nucleotide of the invention, methods of producing the novel polypeptides, antibodies against the polypeptides, methods of detecting the nucleotides or polypeptides in a sample, and methods of identifying compounds which bind to polypeptides of the invention. Although novel, many of the polypeptides of the invention have homology to known proteins, thereby giving an insight into their probable biological activities, and hence potential therapeutic applications. The polypeptides of the invention may have various activities, including cytokine, cell proliferation or cell differentiation activities; stem cell growth factor activity; haematopoiesis regulatory activity; tissue growth activity; immunomodulatory activity; activin- or inhibin-related activities; chemotactic or chemokinetic activities; haemostatic, thrombotic or thrombolytic activities; receptor or ligand activities; or may be involved in oncogenesis, cancer cell proliferation or metastasis. Depending on their biological activities, polypeptides and nucleotides of the invention are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Such conditions include cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell disorders), chronic inflammatory conditions (e.g., asthma or arthritis), proliferative retinopathy, atherosclerosis, coronary heart disease, arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal vascular growth. Polypeptides involved with tissue regeneration and repair (or nucleic acids encoding them) may be used to promote wound healing (e.g., of burns, incisions and ulcers), while those with immunomodulatory activities may be used in the treatment of viral, bacterial and fungal infections in addition to immune disorders. Polypeptides with growth factor activity may be used in cell cultures to promote cell growth. For example, such polypeptides may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease or accidental damage. The polypeptides and nucleotides

CC coded by said sequence belongs to Yap/Rab family and is the homolog of
CC mouse Rab10. The present invention also relates to the polypeptide coded
CC by said nucleotide sequence and the application and preparing process
CC of said polynucleotide and said polypeptide. This sequence encodes a
CC canine Rab10 protein described in the method of the invention.

Sequence 888 BP; 261 A; 202 C; 202 G; 223 T; 0 other;
SQ

Alignment Scores:

Pred. No.:	7.38e-52	Length:	888
Score:	530.50	Matches:	96
Percent Similarity:	82.56%	Conservative:	46
Best Local Similarity:	55.81%	Mismatches:	29
Query Match:	48.01%	Indels:	1
DB:	21	Gaps:	1

US-09-817-198B-2 (1-212) x AAA40109 (1-888)

[illegible]

RESULT 11

AAA96887
ID AAA96887 standard; DNA: 956 BP.

AA
AC
AAA96887:

DT 19-FEB-2001 (first entry)

XX Nucleotide sequence of human RAB10.
DE

XX
KW
RAB protein; GTPase; GTP binding; gene therapy; cancer; ss.

XX
OS Homo sapiens.

XX	Key	Location/Qualifiers
FH	CDS	127..729
FT		/*tag= a
FT		/product= "FAB10"

XX

WO200058464-A2.

104050070M

05-OCT-2000.

1000

13-MAR-2000; 2000WO-USO

25-MAR-1999. 0911S-0126083

22-MAR-1969; 99025

(AXYS-) AXYS PHAF

Allen M, Abel K, McIntosh B, Vega R, Rutter M, Buckler AJ;
WPI: 2000-647233/62.
P-PSDB; AAB19165.

Novel isolated nucleic acid encoding a mammalian RAB protein useful for identifying homologous or related genes, in producing composition that modulates expression or function of RAB for cancer therapy -

Claim 4; Page 38-39; 58pp; English.

The present sequence encodes a mammalian RAB protein. RAB proteins constitute the largest family of small GTPases, with over 40 identified isoforms. RAB proteins contain four highly conserved peptide sequences involved in GTP binding and hydrolysis. Compositions comprising RAB nucleic acid are useful for identifying homologous or related genes, in producing compositions that modulate the expression or function of RAB, for gene therapy, mapping functional regions of the protein and in studying associated physiological pathways. In addition, modulation of the gene activity *in vivo* is used for prophylactic and therapeutic purposes, such as treatment of cancer, and identification of cell type based on expression. The DNA may also be used to identify expression of the gene in a biological specimen.

Sequence 956 BP; 294 A; 205 C; 213 G; 244 T; 0 other;

Alignment Scores:

Pred. No.:	8.16e-52	Length:	956
Score:	530.50	Matches:	96
Percent Similarity:	82.56%	Conservative:	46
Best Local Similarity:	55.81%	Mismatches:	29
Query Match:	48.01%	Indels:	1
DB:	21	Gaps:	1

US-09-817-198B-2 (1-212) X AAA96887 (1-956)

QY 1 MetAlaLysGln--TyrAspValLeuPheArgLeuLeuLeuLeuGlyAspSerGlyVal 19
 Dd 127 ATGGCGAAGAAAGACGTACGACTGCTTTTCAGCTGCTCCTGATCGGGATTCCGGAGTG 18

Qy 20 GlyLysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisIleSer 39
|||||
Dd 187 GGGAGACCTCGTCCTTTTTCGTTTCGGATGATGCCTTCAATACACTATTATTC 24

Qy 40 ThrileGlyValAspPheLysMetLysThrileGluValAspGlyileLysValArgile 59
 |||||:::|||||:::|||||:::|||||::: ||| |||:::|||||:::
 Db 247 ACCATAGGAATAGACTTCAAGATCAAAACAGTTGAATACACGGAAGAGATCAAGCTA 300

Qy 60 GlnIleT⁶⁰TrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArg 79
 |||||
 Db 307 CAGATATGGGATACAGCAGGCCAGGAGCGATTACACCATCACAACTCTACTACAGA 361

Qy 80 ArgAlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIle 99
 ||| ||| ||| ||| ||| ||| ||| |||
Db 367 GGCGCAATGGGTATCATGCTAGTATATGACATCACCATAATGGTAAAAGTTTGAACAATC 423

Qy 100 MetIysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIle 11
 |||||: : : : : |||||: ||| |||: : : : : |||: : : : :
 Db 427 AGCAAAATGGCTTAGAAACATAGATGAGCATGCCAATGAAGATGTGGAAAGAATGTACTA 48

QY	120	GlyAsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlyGlnLeu	13
		: : : : : :	
Db	487	GGAAACAAGTGTGATATGGACGACAAAAGAGTTGTACTTAAGGAAAAAGGAGACAGATT	54

QY 140 AlaLysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLys 159
 Db 547 GCAGGGAGCAGGATAGGTTTTCAGACTAGTCAAAAGCAATATAACATCGAA 606
 QY 160 GluSerPheThrArgLeuThrGluLeuValLeuGln 171
 Db 607 AAGCGTTCCTCACGTTAGCTAGCATATCTTCGA 642

RESULT 12

AAFI16196
 ID AAF16196 standard; cDNA; 1537 BP.

XX AC AAF16196;

XX DT 13-MAR-2001 (first entry)

XX DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:631.

XX KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
 KW neuroprotective; cytoskeletal; cardioactive; immunomodulatory; muscular;
 KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
 KW antibacterial; gene therapy; neural; immune; reproductive; renal;
 KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
 KW wound; infectious disease; ss.

XX OS Homo sapiens.

XX PN WO200055174-A1.

XX PD 21-SEP-2000.

XX PF 08-MAR-2000; 2000WO-US05988.

XX PR 12-MAR-1999; 98US-0124270.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI (ROSE/) ROSEN C A.

XX PI Rosen CA, Ruben SM;

XX DR WPI; 2000-587513/55.

XX DR P-PSDB; AAB56993.

XX PT Prostate cancer associated gene sequences, referred to as prostate
 PT cancer antigens, useful for treatment, prevention, and diagnosis of
 PT disorders such as prostate cancer -

XX PS Claim 1; Page 1073; 2338pp; English.

XX CC AAF15566 to AAF16505 encode the human prostate cancer associated
 CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
 CC The prostate cancer antigens can have neuroprotective, cytoskeletal,
 CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
 CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
 CC and can be used in gene therapy. The prostate cancer antigen
 CC polynucleotides may be used for detection of prostate cancer, chromosome
 CC identification, as chromosome markers, and for numerous other diagnostic
 CC or research purposes. The prostate cancer antigens may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
 CC AAB57303 represent sequences used in the exemplification of the present
 CC invention.

XX SQ Sequence 1537 BP; 405 A; 378 C; 384 G; 369 T; 1 other;

Alignment Scores:

Pred. No.:	1,569-51	Length:	1537
Score:	530.50	Matches:	96
Percent Similarity:	82.56%	Conservative:	46
Best Local Similarity:	55.81%	Mismatches:	29
Query Match:	48.01%	Indels:	1

DB: 21 Gaps: 1

US-09-817-198B-2 (1-212) x AAF16196 (1-1537)

QY 1 MetalAlaLysGln---TyrAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyVal 19
 Db 533 ATGGCGAAGAGACGTCAGCCTGCTTTTCAGCTGCTCCTGATCGGGATTCGGAGTG 592
 QY 20 GlyLysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSer 39
 Db 593 GGGAAAGACCTGCTCTTTTCGTTTTCGGATGATGCTTCAATACCTACCTTTATTC 652
 QY 40 ThrIleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIle 59
 Db 653 ACCATAGGAATAGACTTCAAGATCAAAACAGTTGAATTACAGGAAGAAAGATCAAGCTA 712
 QY 60 GlnIleTyrAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrArg 79
 Db 713 CAGATATGGATACAGCAGGCCAGGAGGATTTTCACACCATCACACCTCTTACTACAGA 772
 QY 80 ArgAlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIle 99
 Db 773 GCGCAATGGGTATCATCTAGTATATGACATCACCAATGGTAAAGTTTGAACATC 832
 QY 100 MetLysTyrValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIle 119
 Db 833 AGCAAAATGGCTTAGAAACATAGATGAGTGCATGAATGATGCGAAAGAAATGTTACTA 892
 QY 120 GlyAsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlyGlnGlnLeu 139
 Db 893 GGAACAAGTGTGATATGGACGACAAAGAGTTGTACCTAAGGAAAGGAGAACAGATT 952
 QY 140 AlaLysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLys 159
 Db 953 GCAAGGGAGCATGGTATTAGGTTTTCAGACTAGTCAAAAGCAATATAACATCGAA 1012
 QY 160 GluSerPheThrArgLeuThrGluLeuValLeuGln 171
 Db 1013 AAGCGTTCCTCACGTTAGCTAGCATATCTTCGA 1048

RESULT 13

AAH17889

ID AAH17889 standard; cDNA; 3533 BP.

XX AC AAH17889;

XX DT 26-JUN-2001 (first entry)

XX DE Human cDNA sequence SEQ ID NO:17618.

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-0116126.

XX PR 29-JUL-1999; 95JP-0248036.

XX PR 27-AUG-1999; 99JP-0300253.

XX PR 11-JAN-2000; 2000JP-0118776.

XX PR 02-MAY-2000; 2000JP-0183767.

XX PR 09-JUN-2000; 2000JP-0241899.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX DR WPI; 2001-318749/34.

Db 438 ATTGACTATGGGATTAAATCTTGGAGACAAGCGCAAAATCCAGTCCAATGTAGAGAG 497

Qy 161 Ser---PheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLys 175

Db 498 GCATTTTTCACACTTCACGAGATATATGACAAAACTCAACAGAAAA 545

Search completed: June 23, 2003, 11:10:07

Job time : 232 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 23, 2003, 10:52:29 ; Search time 1383 Seconds
(without alignments)
2482.605 Million cell updates/sec

Title: US-09-817-198B-2
Perfect score: 1105
Sequence: 1 MAKQYDVLFRLLIGDSGVG.....LEEEKGPEGPANSSKTCWC 212

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1615406 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q/cgn2_1/USPTO.spool/US09817198/runat_18062003_144633_27919/app_query.fasta_1.391
-DB=EST -Qfmt=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=0.1 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcr -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09817198.ecgn_1_1906_erunat_18062003_144633_27919 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hun:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	996	90.1	1091	12	BF160330	BF160330 601768601
2	953.5	86.3	932	12	BF535642	BF535642 602054039
3	922	83.4	616	13	BI648588	BI648588 603277781
4	914	82.7	992	12	BF101730	BF101730 601753464
5	910	82.4	902	12	BF178163	BF178163 601809083
6	885	80.1	699	13	BI557933	BI557933 603237549
7	858.5	77.7	796	12	BF966292	BF966292 602286692
8	838	75.8	1100	12	BF181167	BF181167 601805830
9	829	75.0	616	12	BG247902	BG247902 602359819
10	816	73.8	1121	12	BF182001	BF182001 601805936
11	799	72.3	814	12	BG762967	BG762967 602735022
12	792	71.7	660	13	BG974411	BG974411 602844184
13	777	70.3	911	13	BI649317	BI649317 603278051
14	756	68.4	618	12	BF150465	BF150465 uy86905.y
15	733	66.3	1767	13	BM462147	BM462147 AGENCOURT
16	727	65.8	485	12	BE913096	BE913096 601688580
17	706	63.9	667	13	BI685808	BI685808 603309467
18	704	63.7	843	12	BG865609	BG865609 602783763
19	693.5	62.8	577	12	BG703353	BG703353 602685115
20	686	62.1	465	12	BF565114	BF565114 UT-R-B01-
21	670	60.6	479	12	BF320679	BF320679 uz55d10.y
22	666	60.3	543	12	BG019552	BG019552 daa21e04..
23	663	60.0	539	12	BF022618	BF022618 uy51f07.y
24	658.5	59.6	693	10	BE287683	BE287683 601093322
25	651	58.9	823	13	BM017178	BM017178 603643780
26	649	58.7	385	14	BM936454	BM936454 UT-M-BH2.
27	589	53.3	956	14	BQ944882	BQ944882 AGENCOURT
28	568.5	51.4	704	13	BJ057892	BJ057892 B057692
29	554.5	50.2	714	13	BI259835	BI259835 602971485
30	551.5	49.9	644	14	BQ387194	BQ387194 NISC.mn22
31	547	49.5	611	13	BJ035681	BJ035681 BJ035681
32	547	49.5	678	13	BI395002	BI395002 Bpplp.pk0
33	547	49.5	788	9	AU125167	AU125167 AU125167
34	545	49.3	613	14	BM791355	BM791355 K-EST0071
35	545	49.3	635	9	AU134059	AU134059 AU134059
36	545	49.3	676	9	AU133710	AU133710 AU133710
37	545	49.3	686	12	BG709122	BG709122 602675387
38	545	49.3	816	9	AU119700	AU119700 AU119700
39	545	49.3	835	12	BG425680	BG425680 602452987
40	545	49.3	860	9	AU122740	AU122740 AU122740
41	545	49.3	877	13	BI763900	BI763900 603049666
42	545	49.3	897	14	BQ671010	BQ671010 AGENCOURT
43	545	49.3	986	13	BM475359	BM475359 AGENCOURT
44	545	49.3	1032	13	BM459427	BM459427 AGENCOURT
45	545	49.3	1059	14	BM923211	BM923211 AGENCOURT

ALIGNMENTS

RESULT 1

BF160330

LOCUS

DEFINITION

BF160330

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

BF160330 1091 bp mRNA linear EST 30-OCT-2000
601768601F1 NCI_CGAP_Lu29 Mus musculus cdna clone IMAGE:3987745 5',
mRNA sequence.
BF160330
BF160330.1 GI:11040541
EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1091)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-re@mail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM9177 row: p column: 16
 High quality sequence stop: 706.

FEATURES

source
 1. .992
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:3981183"
 /clone_lib="NCI_CGAP_Maml"
 /tissue_type="tumor, biopsy sample"
 /dev_stage="10 months, virgin"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site.1: SalI;
 Site.2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"
 BASE COUNT 251 a 272 c 306 g 163 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,01e-106 Length: 992
 Score: 914.00 Matches: 181
 Percent Similarity: 86.88% Conservative: 4
 Best Local Similarity: 84.98% Mismatches: 1
 Query Match: 82.71% Indels: 27
 DB: 12 Gaps: 1

US-09-817-198B-2 (1-212) x BF101730 (1-992)

QY 1 MetalLysGlnTyrAspValLeuPheArgLeuLeuLeuLeuLeuLeuValGly 20
 Db 150 ATGGCGMAACAGTACGATGTGCTGTCGGCTACTGCTGATCGGGGACTCCGGGTGGC 209
 QY 21 LysThrCysLeuLeuCysArgPheThrAspGlnPheHisSerSerHisLeuThr 40
 Db 210 AAGACATGCTGCTGTGGCTTCCCGCACACAGAGTTCACATCTCGCATATCCACC 269
 QY 41 IleGlyValAspPheLysMetLysThrIleGlyValAspGlyLysValArgIleGln 60
 Db 270 ATCCGTGTGACTTAAAGTAGAGACTATCGATGTAGACGGCATCAAGTAGAGATACAG 329
 QY 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTrpArg 80
 Db 330 ATTGGGACACAGCAGGGCAGGAGAGTACCAGACTATCACAAAGCAGTACTATCGCGCA 389
 QY 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
 Db 390 GCCCAG-----
 QY 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValcInLysIleLeuIleGly 120
 Db 396 -----TACGCTCCAGAGGAGTCCAGAAAGTCTAATTTGGG 431
 QY 121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnLeuAla 140
 Db 432 AATAAGGCTGATGAAGAGCAGAAACGGCAGGTGGGAGAGCAGGGCAGCAGCTGGCT 491
 QY 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
 Db 492 AAGGAGTAGCGGATGGACTTTCACGAACCAAGTGCCTGCACCACCTTAATATTAAAGAG 551
 QY 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGlnLeuGluLeu 180
 Db 552 TCCTTCACCTGCTGACGGAGGTGGTGTGTCAGGCCCCACAGGAAGAGCTGGATGCTC 611

QY 181 ArgMetArgAlaSerAsnGlnLeuAlaLeuAlaGluLeuGluGluGluGlyLysPyr 200
 Db 612 CGAACACGTGCGCAGCAACGAGCTCGACTGGCCGAGCTGGAGGAGGAGGAGCAACACC 671
 QY 200 cGluGlyProAlaAsnSerSerLysThrCysTrpCys 212
 Db 672 TGAGGGTCCAGCAAACTCTTCAAGACCTGCTGGTGC 708
 RESULT 5
 BF178163 902 bp mRNA linear EST 31-OCT-2000
 LOCUS 601809083F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:4039701 5',
 DEFINITION mRNA sequence.
 ACCESSION BF178163
 VERSION BF178163.1 GI:11056305
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 902)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-re@mail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM9320 row: f column: 22
 High quality sequence start: 2
 High quality sequence stop: 603.

FEATURES

source
 1. .902
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:4039701"
 /clone_lib="NCI_CGAP_Mam5"
 /tissue_type="tumor, gross tissue"
 /dev_stage="7 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site.1: SalI;
 Site.2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigators
 providing samples: Lothar Hennighausen/Robin Humphreys,
 NIH"
 BASE COUNT 216 a 247 c 258 g 181 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 5.63e-106 Length: 902
 Score: 910.00 Matches: 177
 Percent Similarity: 97.28% Conservative: 2
 Best Local Similarity: 96.20% Mismatches: 4
 Query Match: 82.35% Indels: 1
 DB: 12 Gaps: 0

US-09-817-198B-2 (1-212) x BF178163 (1-902)

QY 30 AspaGlnPheHisSerHisIleSerThrIleGlyValAspPheLysMetLysThr 49
 Db 14 GACAACGAGTTCACCTCTCGCATATCTCCACCATCGGTGTGACTTTAAGATGAAGACT 73
 QY 50 IleGluValAspGlyLysValArgIleGlnIleTrpAspThrAlaGlyGlnGluArg 69
 Db 74 ATCGAAGTAGACGGCATCAAGTGAATATACAGATTTGGACACAGCAGGGCAGGAGG 133

Qy 70 TyrGlnThrIleThrLysGlnTyrTyrArgAlaGlnGlnGlyIlePheLeuValTyrAsp 89
 Db 134 TACCAGACTATCAAAAGCAGTACTATCGCGAGCCAGGGAATATTTTAGTCTACGAC 193
 Qy 90 IleSerSerGluArgSerTyrGlnHisIleMetLysTyrValSerAspValAspGluTyr 109
 Db 194 ATTAGCAGTGGCGCTCTATCAGCATATCATGAAGTGGTCACTGAGTGGATGATAC 253
 Qy 110 AlaProGluGlyValGlnLysIleLeuIleGlyAsnLysAlaAspGluGlnLysArg 129
 Db 254 GCTCCAGAAGAGTCCAGAAGATCCTAATTGGGAATAAGCTGATGAAGAGCAAAACGG 313
 Qy 130 GlnValGlyArgGluGlnGlnGlnLeuAlaLys-GluTyrGlyMetAspPheTyrGI 149
 Db 314 CAGGTGGGAGAGAGCAGGAGGAGCAGCTGGTAAACGGATGAGGATGACTTCTACGA 373
 Qy 149 uThrSerAlaCysThrAsnLeuAsnIleLysGluSerPheThrArgLeuThrGluLeuVa 169
 Db 374 AACAAAGTGGCTGCACCAACCTTAATATTAAGAGTCTTTCACCTGCTGACGGAGCTGGT 433
 Qy 169 LeuGlnAlaHisArgLysGluLeuGluGlyLeuArgMetArgAlaSerAsnGluLeuAl 189
 Db 434 GCTCAGGCCACAGGAGCAGAGTGGATGGTCTCCGAACACAGTCCCAACAGGAGCTCGC 493
 Qy 189 aLeuAlaGluLeuGluGlnGluGlyLysProGluGlyProAlaAsnSerSerLysTh 209
 Db 494 ACTGGCCGAGCTGGAGGAGGAGCAAGCAACCTTGAGGGCCAGCAAACTCTTCAAAAGC 553
 Qy 209 rCysTyrCys 212
 Db 554 CTGCTGGTGC 563

RESULT 6
 BI557933
 LOCUS 603237549f1 NCI_GAP_Mam3 Mus musculus cDNA clone IMAGE:5290284 5',
 DEFINITION mRNA sequence.
 ACCSSION BI557933
 VERSION BI557933.1 GI:15445247
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 699)
 TITLE NIH-MGC http://mgi.nci.nih.gov/.
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cchapbs-r@mail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
 cDNA Library Prepared by: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM11734 row: b column: 13
 High quality sequence stop: 698.
 Location/Qualifiers
 1..699
 /organism="Mus musculus"
 /strain="129,C57BL/6J,FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5290284"
 /clone_lib="NCI_GAP_Mam3"
 /tissue_type="tumor, gross tissue"
 /dev_stage="10 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;
 Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Library constructed by Life Technologies. Investigators
 providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH

FEATURES
 source
 1..699
 /organism="Mus musculus"
 /strain="129,C57BL/6J,FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5290284"
 /clone_lib="NCI_GAP_Mam3"
 /tissue_type="tumor, gross tissue"
 /dev_stage="10 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;
 Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Library constructed by Life Technologies. Investigators
 providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH

Reference for transgenic model: Xu et al., Nature Genetics
 22, 37-43 (1999).
 BASE COUNT 164 a 188 c 219 g 128 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 6.le-103 Length: 699
 Score: 885.00 Matches: 172
 Percent Similarity: 98.87% Conservative: 3
 Best Local Similarity: 97.18% Mismatches: 1
 Query Match: 80.09% Indels: 1
 Db: 13 Gaps: 0

US-09-817-198B-2 (1-212) x BI557933 (1-699)

Qy 1 MetaLysGlnTyrAspValLeuPheArgLeuLeuIleGlyAspSerGlyValGly 20
 Db 156 ATGGCGAAGACAGTACGATGCTGTCGGGTACTGCTGATCGGGGACTCGGGGTGGC 215
 Qy 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSerThr 40
 Db 216 AAGACATGCTGCTGTGCGCTTCCGACAAACAGTTCACACTCTCGCATATCTCCACC 275
 Qy 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60
 Db 276 ATCGGTGTGACTTAAAGATGAAGACTATCATGATAGACGGCATCAAAGTGAGATACAG 335
 Qy 61 IleThrAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
 Db 336 ATTTGGGACACAGCAGGCGAGGAGGTACCACTATCACAAAGCAGTACTATTCGGCGA 395
 Qy 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
 Db 396 GCCCAGGGAATATTTTAGTCTACGACATTAAGCAGTGGAGCGCTCTATCAGCATATCATG 455
 Qy 101 LysThrValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleGly 120
 Db 456 AAGTGGGTCACTGAGTGGATGATGATGCTCCAGAAAGAGTCCAGAAAGATCTTAATTTGGG 515
 Qy 121 AsnLysAlaAspGluGlnLysArgGlnVal-GlyArgGluGlnGlnLysLeuAl 140
 Db 516 AATCAGGCTGATGAAGACAGCAAGCGGAGTGGGCGAGAGAGCGGGGAGCAGCTGGC 575
 Qy 140 aLysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGI 160
 Db 576 TAAGAGTACGGCATGGACTTCTACGAACAAGTGCCTGCACCAACCTTAATATTAAGA 635
 Qy 160 uSerPheThrArgLeuThrGluLeuValLeuAlaHisArgLysGlu 176
 Db 636 GTCTTCACTGCTGACGGAGTGGTCTGCGAGGCCGCCAGGACACAG 684

RESULT 7
 BF966292
 LOCUS 602286692F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4375366 5',
 DEFINITION mRNA sequence.
 ACCSSION BF966292
 VERSION BF966292.1 GI:12333507
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 796)
 TITLE NIH-MGC http://mgi.nci.nih.gov/.
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cchapbs-r@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10040 row: h column: 23
High quality sequence stop: 663.

FEATURES	SOURCE
-----------------	---------------

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4375366"
/clone_lib="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); site_1: BamHI; Site_2: SalI-XhoI (ctcag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.5 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
202 a 1393 c 263 t 137 t 1 others

```

Alignment Scores:		
Pred. No.:	1.89e-99	Length:
Score:	856.50	Matches:
Percent Similarity:	75.77%	Conservative:
Best Local Similarity:	75.00%	Mismatches:
Query Match:	77.69%	Indels:
DB:	12	Gaps:
		1
US-09-817-198B-2 (1-212)	x	BF966292 (1-796)

```

QY 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisLeuSerThr 40
D 220 AAGACATGCGTGTGCGCTTCACCGACAGAGTTCACCTCTCGCATATCTCCACC 279
QY 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60
D 280 ATCGGTGTTGACCTTAAGATGAAGACTATCGATGATAGCGGCATCAAGAGAGAAATACAG 339
QY 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrArgArg 80
D 340 ATTTGGGACACAGCGGAGAGAGTACACACTATCACAAAGCAGTACTATCGCGGA 399
QY 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
D 400 GCCCAGGGAATATTTAGTCTACGACATTAGCAGTACGCGCTCTATCAGCATATCATG 459
QY 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuGly 120
D 460 AAGTGGTCACTGACGTGGATGATGAGTCCAGAGAGGATCCAGAAATCTTAATTTGGG 519
QY 121 AsnLysAlaAspGluGlnGlnLysArgGlnVal-GlyArgGluGlnGlnGlnLeuAl 140
D 520 AATAAGGCTGATCAGACGACAGAACGCGCAGGTGGGAGCAGACGCGGCGCAGCTGGC 579
QY 140 aLysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsn--IleLys 159
D 580 TAAGGAGTACGGATGACTTCTACGAAACAAAGTGGCTGGAACAACTTAACTATTACAC 639
QY 160 GluSerPheThrArg-LeuThrGluLeuValLeuGlnAlaHisArgLys-GluLeuGluG 179
D 640 GATCCTTCACTCGCTTGACGAGCTGGTCTGCGAGGCCACAGGAAGCAGCTGGACT 699
QY 179 lYLeuArgMetArgAlaSerAsnGluLeu-----AlaLeuAlaGluL 193
D 700 GG-----GCTCCGACACAGCGTGGCCACAGCAGCGCGCGCTGGCGAAC 747
QY 193 eu-----GluGluGluGluGlyLysProGluGly-----ProAlaAsnSers 207
D 748 TTGGGAGGAGCGGACGACAGGAGAAACCCGAGGGGCGCGGACAAACCTTCCCAAGACA 807
QY 207 er 207
D 808 CC 809

RESULT 9
BG247902
LOCUS
DEFINITION
602359819F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4488197 5',
mRNA sequence.
ACCESSION
BG247902
VERSION
BG247902.1 GI:12757717
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS
NIH-MGC http://mgi.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10334 row: f column: 06
High quality sequence stop: 616.
Location/Qualifiers
FEATURES

```

source

```

1. 616
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4488197"
/clone_lib="NCI_CGAP_Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"

```

```

/note="Organ: mammary; Vector: pCMV-SPORT6; Site.1: SalI;
Site.2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

```

```

BASE COUNT      165 a      152 c      183 g      116 t
ORIGIN

```

Alignment Scores:

```

Pred. No.:      7.73e-96      Length:      616
Score:          829.00      Matches:      158
Percent Similarity: 99.38%      Conservative: 3
Best Local Similarity: 97.53%      Mismatches: 1
Query Match:      75.02%      Indels:      0
DB:              12      Gaps:      0

```

```

US-09-817-198B-2 (1-212) x BG247902 (1-616)

```

```

QY 51 GluValAspGlyIleLysValArgIleGlnIleTrpAspThrAlaGlyGlnGluArgTyr 70
D 3 GATGTAGACGGCATCAAGTGAGATACAGATTGGGACACAGCAGGCGAGGAGGTAC 62
QY 71 GlnThrIleThrLysGlnTyrTyrArgArgAlaGlnGlyIlePheLeuValTyrAspIle 90
D 63 CAGACTATCAAAAGCAGTACTATCGCGGAGCCAGGGAATATTTTGTCTACGACATT 122
QY 91 SerSerGluArgSerTyrGlnHisIleMetLysTrpValSerAspValAspGluTyrAla 110
D 123 AGCAGTGGCGCTCTATCAGCATATCATGAAGTGGTCACTGAGTGGATGATGATGCT 182
QY 111 ProGluGlyValGlnLysIleLeuIleGlyAsnLysAlaAspGluGlnLysArgGln 130
D 183 CCAGAGGAGTCCAGAAAGATCCTAATTGGGAATAAGCTGATGAAGACGAAACGGCAG 242
QY 131 ValGlyArgGluGlnGlyGlnGlnLeuAlaLysGluTyrGlyMetAspPheTyrGluThr 150
D 243 GTGGGAGAGAGCAGGCGGCGCAGCTGGTGGTTCGAGTACGCGATGGACTTCTACGAAACA 302
QY 151 SerAlaCysThrAsnLeuAsnIleLysGluSerPheThrArgLeuThrGluLeuValLeu 170
D 303 AGTGCCCTGCACCAACCTTAATATTAAGAGTCTCTACTCTGTCGCGGAGCTGGTCTG 362
QY 171 GlnAlaHisArgLysGluLeuGluGlyLeuArgMetArgAlaSerAsnGluLeuAlaLeu 190
D 363 CAGGCCACAGGAAAGAGCTGGTGGTCTCCGAACAGCTGCCACACAGCTGCCACTG 422
QY 191 AlaGluLeuGluGluGluGluGlyLysProGluGlyProAlaAsnSerSerLysThrCys 210
D 423 GCCGAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 482
QY 211 TrpCys 212
D 483 TGTGTGC 488

```

RESULT 10

```

BF182001
LOCUS
DEFINITION
601805936F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:4036664 5',
mRNA sequence.
ACCESSION
BF182001
VERSION
BF182001.1 GI:11060143
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```



```

|||||
397 GCCAGGGGAATATTTAGTCTTAGACATTTAGCAGTGGCTCCCTATCAGCATATCATG 456
|||||
101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
|||||
457 AAGTGGGTGAGTGGAGTGGATGAGTGGCTCCAGAGAGAGTCCAGAGATCCCTAATTGGG 516
|||||
121 AsnLysAlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAla 140
|||||
517 AATAATGCTGATGAAGAGCAGAAACCGCAGGTGGGAGAGAGAGAGTGCAGCAGCTGGC 576
|||||
141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThr 154
|||||
577 TTAGAGTACGGCAGTGGACTTCTAGAAACAAAGTGCCTGCACC 618
|||||

RESULT 15
BM462147
LOCUS
DEFINITION BM462147 1767 bp mRNA linear EST 05-FEB-2002
5', mRNA sequence.
ACCESSION BM462147
VERSION BM462147.1 GI:18511187
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo..
REFERENCE
1 (bases 1 to 1767)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM12111 row: g column: 19
High quality sequence start: 89
High quality sequence stop: 574.
Location/Qualifiers
1. 1767
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:5491242"
/clone_lib="NIH_MGC_67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
BASE COUNT 318 a 704 c 411 g 320 t 14 others
ORIGIN

Alignment Scores:
Pred. No.: 8.62e-83 Length: 1767
Score: 733.00 Matches: 141
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.33% Indels: 0
DB: 13 Gaps: 0

US-09-817-198B-2 (1-212) x BM462147 (1-1767)

QY 72 ThrIleThrLysGlnTyrTyrArgAlaGlnGlyIlePheLeuValTyrAspIleSer 91
|||||
DB 96 ACCATCAACAAGCAGTACTATCGCGGCCCGCCAGGGGATATTTGGTCTATGACATTAGC 155
|||||
QY 92 SerGluArgSerTyrGlnHisIleMetLysTrpValSerAspValAspGluTyrAlaPro 111
|||||

```

```

|||||
156 AGCAGGCGCTTTACCAGCACATCATGAAGTGGTCACTGACGTGGATGATAGTCACCA 215
|||||
112 GluGlyValGlnLysIleLeuIleGlyAsnLysAlaAspGluGluGlnLysArgGlnVal 131
|||||
216 GAAGGCGTCCAGAGATCCCTTATTTGGGAATAAGCTCATGAGGAGCAGAAACGCGAGTG 275
|||||
132 GlyArgGluGlnGlyGlnGlnLeuAlaLysGluTyrGlyMetAspPheTyrGluThrSer 151
|||||
276 GGAAGAGACGAAGGGCAGCAGCTGGCGCAAGAGTATGGCATGGACTTCTATGAACAAGT 335
|||||
152 AlaCysThrAsnLeuAsnIleLysGluSerPheThrArgLeuThrGluLeuValLeuGln 171
|||||
336 GCCTGCACCAACCTCAACATTAAGAGTCTTACGCGTCTGACAGAGCTGGTCTGCAG 395
|||||
172 AlaHisArgLysGluLeuGluGlyLeuArgMetArgAlaSerAsnGluLeuAlaLeuAla 191
|||||
396 GCCCATAGGAAGGAGCTGGGAAGGCTCCGGATGCGTCCAGCATGATGATGGCACTGGCA 455
|||||
192 GluLeuGluGluGluGluGlyLysProGluGlyProAlaAsnSerSerLysThrCysTrp 211
|||||
456 GAGCTGGAGGAGGAGGAGGAGGCAACCCGAGGGCCCGCAACTCTTCGAAAACCTGCTGG 515
|||||
QY 212 Cys 212
|||||
DB 516 TGC 518
|||||

```

Search completed: June 23, 2003, 12:15:52
Job time : 1390 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 23, 2003, 08:15:02 ; Search time 42 seconds
(without alignments)
1547.987 Million cell updates/sec

Title: us-09-817-198b-2

Perfect score: 1105

Sequence: 1 MAKQDVLFRLILLIGDSGVG.....LEEEGKPGPANSSKTCWC 212

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-frame+ p2n.model
-Q/cgn2_1/USPTO_spool/US09817198/runat_18062003_144633_27928/app_query.fasta_1.391
-DB-Issued Patents_NA -QFMT-fastap -SUFFIX-p2n.rni -MINMATCH-0.1 -LOOPCL-0
-LOOPEXT-0 -UNITS-bits -START-1 -END-1 -MATRIX-blosum62 -TRANS-human40.cdi
-LIST-45 -DOALIGN-200 -THR SCORE-pct -THR_MAX-100 -THR_MIN-0 ALIGN-15
-MODE-LOCAL -OUTFWT-pto -NORM-ext -HEAPSIZE-500 -MINLEN-0 -MAXLEN-2000000000
-USER-US09817198.ecgn_1_1_40/runat_18062003_144633_27928 -NCPU-6 -ICPU-3
-NO_MMAP -LARGEQUERY -NEG_SCORES-0 -WAIT -DSPBLOCK-100 -LONGLOG
-DEV_TIMEOUT-120 -WARN_TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPOP-0.5 -FGAPOPOP-6
-FGAPOEXT-7 -YGAPOP-10 -YGAPOP-0.5 -DELOP-6 -DELEXT-7

Database : Issued Patents_NA:**

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	478.5	43.3	925	2	US-08-916-901-4
2	478.5	43.3	925	4	US-09-154-602-4
3	437.5	39.6	1340	2	US-08-824-873-2
4	437.5	39.6	1340	3	US-09-198-184-2
5	429	38.8	875	4	US-09-075-454-10
6	429	38.8	2612	4	US-09-484-9705-142
7	419	37.9	639	4	US-09-399-913-66
8	411.5	37.2	970	3	US-08-888-077A-28
9	373	33.8	847	2	US-08-773-423-4
10	343	31.0	890	3	US-08-741-411-4
11	340.5	30.8	803	4	US-09-075-454-13
12	321.5	29.1	1407	4	US-09-493-914-1
					Sequence 4, Appl1
					Sequence 4, Appl1
					Sequence 2, Appl1
					Sequence 10, Appl1
					Sequence 142, App
					Sequence 66, Appl
					Sequence 28, Appl
					Sequence 4, Appl1
					Sequence 4, Appl1
					Sequence 13, Appl1
					Sequence 1, Appl1

13	320	29.0	1172	4	US-09-075-454-8	Sequence 8, Appl1
14	320	29.0	1533	4	US-09-075-454-11	Sequence 11, Appl1
15	313	28.3	1255	2	US-08-766-551-6	Sequence 6, Appl1
16	297	26.9	1175	2	US-08-773-423-6	Sequence 2, Appl1
17	282	25.5	848	3	US-08-741-411-2	Sequence 2, Appl1
18	274.5	24.8	1749	4	US-09-149-476-54	Sequence 54, Appl1
19	271.5	24.6	820	3	US-08-741-411-6	Sequence 6, Appl1
20	262	23.7	615	1	US-08-247-946A-5	Sequence 5, Appl1
21	262	23.7	615	5	PCT-US95-06420-5	Sequence 5, Appl1
22	260	23.5	985	4	US-08-842-306B-1	Sequence 1, Appl1
23	260	23.5	985	4	US-08-838-973B-1	Sequence 1, Appl1
24	260	23.5	985	4	US-08-771-212A-1	Sequence 1, Appl1
25	259.5	23.5	3198	4	US-08-842-306B-48	Sequence 48, Appl1
26	259.5	23.5	3198	4	US-08-838-973B-48	Sequence 48, Appl1
27	255	23.1	600	3	US-09-078-317-1	Sequence 1, Appl1
28	255	23.1	600	4	US-09-454-818-1	Sequence 1, Appl1
29	255	23.1	2309	3	US-09-078-317-3	Sequence 3, Appl1
30	255	23.1	2309	4	US-09-454-818-3	Sequence 3, Appl1
31	249.5	22.6	624	4	US-09-415-522-1	Sequence 1, Appl1
32	249.5	22.6	5197	4	US-09-293-170-6	Sequence 6, Appl1
33	246	22.3	1074	2	US-09-156-424-1	Sequence 1, Appl1
34	246	22.3	1074	4	US-09-387-341-1	Sequence 1, Appl1
35	245.5	22.2	603	4	US-09-325-932A-29	Sequence 29, Appl1
36	245.5	22.2	932	4	US-09-325-932A-28	Sequence 28, Appl1
37	245.5	22.2	2436	1	US-08-306-891B-16	Sequence 16, Appl1
38	244.5	22.1	574	2	US-08-429-964-83	Sequence 83, Appl1
39	239	21.6	607	2	US-08-429-964-85	Sequence 85, Appl1
40	236	21.4	1058	3	US-09-156-807-1	Sequence 1, Appl1
41	236	21.4	5775	1	US-08-306-691B-15	Sequence 15, Appl1
42	236	21.4	5775	5	PCT-US93-06251-29	Sequence 29, Appl1
43	232	21.0	1284	2	US-09-161-015-1	Sequence 1, Appl1
44	232	21.0	1284	4	US-09-387-341-150	Sequence 150, App
45	230.5	20.9	1166	5	PCT-US96-12129B-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1

US-08-916-901-4
; Sequence 4, Application US/08916901
; Patent No. 5892012
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: RAB PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/916,901
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0367 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166

QY 121 AsnLysAlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAla 140
 Db 426 AACAGAGCGGCTTACCACCAAGAGGTGGTGGACCAACACACAGCAAGAGTTTGA 485
 QY 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
 Db 486 GACTCTCGGCATCCCTTCTTGGACAGCAGCCCAAGATGCCCAATGTGGACGAG 545
 QY 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgTysGluLeuGluGlyLeu 180
 Db 546 GCGTTC-----ATGACCATGCTCTGAAATCAAAAGCGGATGGGCGCT 590
 QY 181 ArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGluGlyLysPro 200
 Db 591 GGAGCAGCGCTCTGGGGGCGAGCGGCCCAATCTCAAGATCGACACGCCCTGTAAAGCGG 650
 QY 201 Glugly 202
 Db 651 GCTGGC 656

RESULT 3
 US-08-824-873-2
 ; Sequence 2, Application US/08824873
 ; Patent No. 5843717
 ; GENERAL INFORMATION:
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Guegler, Karl
 ; TITLE OF INVENTION: NOVEL RAB PROTEIN
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/824,873
 ; FILING DATE: Filed Herewith
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0240 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1340 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: FANCMOT04
 ; CLONE: 738957
 US-08-824-873-2

Alignment Scores: 6.5e-46 Length: 1340
 Pred. No.: 90
 Score: 437.50 Matches: 90
 Percent Similarity: 67.84% Conservative: 45
 Best Local Similarity: 45.23% Mismatches: 52
 Query Match: 39.59% Indels: 13
 Db: 2 Gaps: 3

US-09-817-198B-2 (1-212) x US-08-824-873-2 (1-1340)
 QY 5 TyrAspValLeuPheArgLeuLeuIleGlyAspSerGlyValGlyLysThrCysLeu 24
 Db 22 TACACGCTCGCTTCAAGGTCTGCTGGTGGGAGACTCGGTGTGGGAGAACACTGTCTG 81
 QY 25 Leu-CysArgPheThrAspAsnGluPhe---HisSerSerHisIleSerThrIleGlyVa 43
 Db 82 CTGGGTGCGATTTCAAGGATGGTGTCTTCCCTGGCGGGGACCTTCATCTCCACCGTAGC-AT 140
 QY 43 lAspPheLysMetLysThrIleGluValAlaGlyLysValArgIleGlnIleTrpAs 63
 Db 141 TGACTTCGGGAACAAAGTTCTGGACGTGGATGGTGAAGTGAAGTGCAGATGTGGGA 200
 QY 63 pThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgAlaGlnGln 83
 Db 201 CACAGCTGGTCAGGAGCGGTTCCGACGTGTACCCATGCCTACTACCGGGATGCTCATGC 260
 QY 83 yIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMetLysTrpVa 103
 Db 261 TCTGCTGCTGCTTACGATGTCAACACAGGCGCTCTTTGACACATCCAGGCGCTGCT 320
 QY 103 lSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGlyAsnLysAl 123
 Db 321 GACCGAGATCCACGAGTACGCCACGACGAGCGTGGCGCTCATGCTGTGGGGAACAAGGT 380
 QY 123 aAspGluGlnGlnLysArgGlnValGlyArgGluGlnGlnGlnGlnLeuAlaLysGluTy 143
 Db 381 GGACTCTGCCCATGAGCGTGTGGTGAAGAGGAGGACGGGAGAGAGCTGGCCNAGAGTA 440
 QY 143 rGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSerPheTh 163
 Db 441 TGGACTGCCCTTCATGGAGACGAGCGGCCCAAGACGCGCTCAACGTGGAGCTTGGCCTT 500
 QY 163 rArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeuArgMetAr 183
 Db 501 AGCCATAGCA-----AAGAGTGAAGCAGCGCTCCATGAA 536
 QY 183 gAlaSerAsnGlu-----LeuAlaLeuAlaGluLeuGluGluGluGlyLys 199
 Db 537 GGCTCCCGAGCGCGCGCTTCCGGCTGCATGATTACGTTAAGAGGGAGGTCGA 591

RESULT 4
 US-09-198-184-2
 ; Sequence 2, Application US/09198184
 ; Patent No. 6010859
 ; GENERAL INFORMATION:
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Guegler, Karl
 ; TITLE OF INVENTION: NOVEL RAB PROTEIN
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/198,184
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/824,873
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.


```

Db 312 TTGCTTCTGCTGATGACATCACCACAAATCTTCTTCGACAAACATCAGGGCTGGCTC 371
QY 104 SerAspValaspGluTyrAlaProGluGlyValGlnLysIleLeuIleGlyAsnLysAla 123
Db 372 ACTGAGATTATCATGATGATGCCAGAGGGAGCTGGTGTATCATGCTGTAGGCAACAGGCG 431
QY 124 AspGluGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLysGluTyr 143
Db 432 GATATGAGCAGCGAAGAGTGTATCGGTTCGGAAGACGGAGACCTTGGCCAGGAGTAC 491
QY 144 GlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSerPheThr 163
Db 492 GGTGTTCCTTCTCTGGAGACCGCCAGACACTGGCATGAATGTG----- 536
QY 164 ArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeuArgMetArg 183
Db 537 -----GAGTTAGCTTCTTGGCCATCGCCCAAGGAA-----CTGAATAACCGG 578
QY 184 AlaSerAsnGlu 187
Db 579 GCCGGGCATCAG 590

```

RESULT 6

```

US-09-484-970B-142
; Sequence 142, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmut, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 142
; LENGTH: 2612
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6426186 412477.1CB1
US-09-484-970B-142

```

```

Alignment Scores:
Pred. No.: 2,08e-44 Length: 2612
Score: 429.00 Matches: 82
Percent Similarity: 68.48% Conservativity: 44
Best Local Similarity: 44.57% Mismatches: 46
Query Match: 38.82% Indels: 12
DB: 4 Gaps: 3

```

US-09-817-198B-2 (1-212) x US-09-484-970B-142 (1-2612)

```

QY 5 TyrAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyValGlyLysThrCysLeu 24
Db 91 TAGCGACTCAGCGGCAAGGTGATGCTCTGGGAGACACAGGCGTCGCAAAACATGTTTC 150
QY 25 LeuCysArgPheThrAspAsnGluPheHisSer---SerHisIleSerThrIleGlyVal 43
Db 151 CTGATCCATTTAAAGACGGGGCTTCCGTTCGCGAACCTTCATAGCCACCGTCGCGCAT 210
QY 44 AspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGlnIleTyrAsp 63
Db 211 GACTTCAGGACAAAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 270
QY 64 ThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArgAlaGlnGly 83
Db 271 ACCGCTGGGAGGACGGTTCGAGAGCTCACCCTGCTATTACAGAGATGCTCAGGCC 330
QY 84 IlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMetLysTrpVal 103

```

```

Db 331 TTGCTTCTGCTGATGACATCACCACAAATCTTCTTCGACAAACATCAGGGCTGGCTC 390
QY 104 SerAspValaspGluTyrAlaProGluGlyValGlnLysIleLeuIleGlyAsnLysAla 123
Db 391 ACTGAGATTATCATGATGATGCCAGAGGGAGCTGGTGTATCATGCTGTAGGCAACAGGCG 450
QY 124 AspGluGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLysGluTyr 143
Db 451 GATATGAGCAGCGAAGAGTGTATCGGTTCGGAAGACGGAGAGACCTTGGCCAGGAGTAC 510
QY 144 GlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSerPheThr 163
Db 511 GGTGTTCCTTCTCTGGAGACCGCCAGACACTGGCATGAATGTG----- 555
QY 164 ArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeuArgMetArg 183
Db 556 -----GAGTTAGCTTCTTGGCCATCGCCCAAGGAA-----CTGAATAACCGG 597
QY 184 AlaSerAsnGlu 187
Db 598 GCCGGGCATCAG 609

```

RESULT 7

```

US-09-399-913-66
; Sequence 66, Application US/09399913
; Patent No. 6361971
; GENERAL INFORMATION:
; APPLICANT: Rhodes, Kenneth
; APPLICANT: Betty, Maria
; APPLICANT: Ling, Hual-Ping
; APPLICANT: An, Wenqian
; TITLE OF INVENTION: POTASSIUM CHANNEL INTERACTORS AND USES THEREFOR
; FILE REFERENCE: MNI-070CP2
; CURRENT APPLICATION NUMBER: US/09/399,913
; CURRENT FILING DATE: 1999-09-21
; EARLIER APPLICATION NUMBER: USSN 60/110,277
; EARLIER FILING DATE: 1998-11-30
; EARLIER APPLICATION NUMBER: USSN 60/110,033
; EARLIER FILING DATE: 1998-11-25
; EARLIER APPLICATION NUMBER: USSN 60/109,333
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: USSN 09/298,731
; EARLIER FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: USSN 09/350,614
; EARLIER FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: USSN 09/350,874
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 66
; LENGTH: 639
; TYPE: DNA
; ORGANISM: Rattus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(636)
US-09-399-913-66

```

```

Alignment Scores:
Pred. No.: 4.8e-44 Length: 639
Score: 419.00 Matches: 85
Percent Similarity: 64.59% Conservativity: 50
Best Local Similarity: 40.67% Mismatches: 65
Query Match: 37.92% Indels: 9
DB: 4 Gaps: 2

```

US-09-817-198B-2 (1-212) x US-09-399-913-66 (1-639)

```

QY 5 TyrAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyValGlyLysThrCysLeu 24
Db 7 TAGCCCTATCTCTTCAAGTACATCATCGGCGACACAGGTGTTGTTAAATCGTCTTA 66

```

```

QY 25 LeuCyAspGpHeThrAspAsnGluPheHisSerHisIleSerThrIleGlyValAsp 44
Db 67 TTGCTACATTTACAGACAAGGTTTCAGCGGTGCATGACCTCACATTTGGTGTAG 126
QY 45 PheLysMetLysThrIleGluValAspGlyIleLysValArgIleGlnIleThrAspThr 64
Db 127 TTTGCTGCTCGAATGATAACCATTCATGGGAACAGATAAATCCAGATCTGGGATACA 186
QY 65 AlaGlyGlnGluArgTyGlnThrIleThrLysGlnTyArgArgAlaGlnGlyIle 84
Db 187 GCAGGGCAGGAGTCTTCGTTCATCAACAAGCATATACAGAGGTGCAGCGGGGCT 246
QY 85 PheLeuValTyAspIleSerSerGluArgSerTyGlnHisIleMetLysTrpValSer 104
Db 247 TTACTAGTGTATGATATTACAGAGAGACACGTTCAACCACTTGACAACTGGTAGAA 306
QY 105 AspValAspGluTyAlaProGluGlyValGlnLysIleLeuIleGlyAsnLysAlaAsp 124
Db 307 GACGCCGTCAGCATTCCTCAATTCACCATGTCATCATGCTTATTTGGAATAAAGTGAC 366
QY 125 GluGluGlnLysArgGlnValGlyArgGlnGlnGlnLysAlaLysGluTyGly 144
Db 367 TTAGATCTAGGAGAGAGTGAAGAAAGGAAGAGTGAAGCTTTTGCAGAGAGCATGA 426
QY 145 MetAspPheTyGluThrSerAlaCysThrAsnLeuAsnIleLysGluSerPheThrArg 164
Db 427 CTTATCTTCATGGAATCTTCCCAAGACTGCTCTTAATGTAGAGGAGCATTTATTAA 486
QY 165 LeuThrGluLeuVal-LeuGlnAlaHisArgLysGluLeuGluGlyLeuArgMetArg-- 183
Db 487 ACAGCAAAAGAAATTTATGAAATAATCCAAAGAGGGTCTTTGACATTAAATAGAGCA 546
QY 184 -AlaSerAsnGluLeuAlaLeuAlaGluLeu-----GluGluGlu 196
Db 547 RACGGCATCAATATGGCCCTCAGCATGCTGTACCAATGCATCTCAGGGAGCAACCA 606
QY 196 uGluGlyLysProGluGlyProAla 204
Db 607 GGAGGCGCAGCAGGAGGGGAGGCT 631

RESULT 8
US-08-888-077A-28
; Sequence 28, Application US/08888077A
; Patent No. 6020143
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE AND USES THEREFOR.
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LERNER, DAVID, LITTENBERG, KRUMHOLZ & MENTLIK
; STREET: 600 SOUTH AVENUE WEST
; CITY: WESTFIELD
; STATE: NJ
; COUNTRY: USA
; ZIP: 07090-1497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/888,077A
; FILING DATE: 03-JUL-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/592,541
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: PALISI, THOMAS M
; REGISTRATION NUMBER: 36,629

;
; REFERENCE/DOCKET NUMBER: SCHERING 3.0-017 CIP CIP CIP IV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 654-5000
; TELEFAX: (908) 654-7866
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 970 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..970
; OTHER INFORMATION: /note= "Y2H9"
US-08-888-077A-28

Alignment Scores: 7.94e-43 Length: 970
Pred. No.: 411.50 Matches: 85
Score: 64.04% Conservative: 45
Percent Similarity: 41.87% Mismatches: 58
Best Local Similarity: 37.24% Indels: 5
Query Match: 3 Gaps: 3
DB:

US-09-817-198B-2 (1-212) x US-08-888-077A-28 (1-970)
QY 4 GlnTyAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyValGlyLysThrCys 23
Db 72 GAGTACGACTACCTCTTTAAAGTTGCTCTTATTGAGATTCTGCTTTGGAAGAGTAAT 131
QY 24 LeuLeuCysArgPheThrAspAsnGluPheHisSerHisIleSerThrIleGlyVal 43
Db 132 CTCCTGCTCGATTACTCGAAATGAGTTTAACTCGAAACCAAGACACCATTTGGAGTA 191
QY 44 AspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGlnIleTrpAsp 63
Db 192 GAGTTTCCACAAGACATCCAGTTGATCGAAACCAATTAAGGCACAGATATGGAC 251
QY 64 ThrAlaGlyGlnGluArgTyGlnThrIleThrLysGlnTyArgArgAlaGlnGly 83
Db 252 ACAGCAGGCGAAGACGATATCGAGCTATAACATCATATTTATCTGCGAGCTGTAGT 311
QY 84 IlePheLeuValTyAspIleSerSerGluArgSerTyGlnHisIleMetLysTrpVal 103
Db 312 GCCTTATTGGTTTATGACATTTGCTTAACATCTCAGATTAAGAAATGTAGACGGTGTG 371
QY 104 SerAspValAspGluTyAlaProGluGlyValGlnLysIleLeuIleGlyAsnLysAla 123
Db 372 AAAGAAGCTGAGAGATCATGCTGATAGTAACATTTGTTATCTGCTTGTGGCAATAAGAGT 431
QY 124 AspGluGlnGlnLysArgGlnValGlyArgGlnGlnGlnGlnLysGluTy 143
Db 432 GATCTACGTCATCTCAGGGCAGTCTCTACAGATGAAGCAAGAGCTTTTGCAGAAAGAA 491
QY 144 GlyMetAspPheTyGluThrSerAlaCysThrAsnLeuAsnIleLysGluSerPhe-- 162
Db 492 GGTTCATTCATTTGAAGATTCGCGCCCTAGACTCTACAATGTAGAGCTGCTTTTCAG 551
QY 163 ThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeuArg-- 181
Db 552 ACAATTTTAAACAGAGATTTTACCGCATTTCTTCTCAGAAAGCAAAATCTCAGACAGAC 611
QY 182 -----MetArgAlaSerAsnGluLeuAlaLeuAlaGluLeu---GluGluGluGly 198
Db 612 ATGACATGCTCTCCAAGCAACAATGTGTTCTCTATTTATGTTTCCACCACCACTGANAAC 671
QY 199 LysProGlu 201
Db 672 AACCCAAAG 680

RESULT 9
US-08-773-423-4
; Sequence 4, Application US/08773423
```



```
QY 142 uTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSerPh 162
||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 486 GTATGGGCAATTATTCTGTGAACAAGTGCACAAAGATGGTTCTAACATAGATGGAGGCTGT 545
QY 162 eThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGlu----- 176
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 546 TCTGCACCTTGCCTCGAAGTGAAGAGAACTGCACAGGATGCACAGACATCCATTAC 605
QY 177 ----LeuGluGlyLeuArgMetArgAlaSerAsnGluLeu 188
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 606 CAATCTAACCGGAGCAATTCACAAAGATCACCACAGATG 645

RESULT 12
US-09-493-914-1
; Sequence 1, Application US/09493914
; Patent No. 6448073
; GENERAL INFORMATION:
; APPLICANT: Jger, Dick
; APPLICANT: Stockert, Elizabeth
; APPLICANT: Jger, Elke
; APPLICANT: Gure, Ali
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Kuth, Alexander
; APPLICANT: Old, Lloyd J.
; APPLICANT: Chen, Yao-Tseng
; TITLE OF INVENTION: Cancer Associated Antigen Encoding Nucleic Acid Molecules
; FILE REFERENCE: LUD-5638
; CURRENT APPLICATION NUMBER: US/09/493,914
; CURRENT FILING DATE: 2000-01-28
; NUMBER OF SEQ ID NOS: 5
; SEQ ID NO 1
; LENGTH: 1407
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: 1165 .. 1390
US-09-493-914-1

Alignment Scores:
Pred. No.: 3-5e-31 Length: 1407
Score: 321.50 Matches: 83
Percent Similarity: 54.19% Conservative: 40
Best Local Similarity: 36.56% Mismatches: 66
Query Match: 29.10% Indels: 38
DB: 4 Gaps: 10

US-09-817-198B-2 (1-212) x US-09-493-914-1 (1-1407)
QY 8 LeuPheArgLeuLeuLeuIleGlyAspSerGlyValGlyLysThrCysLeuLeuCysArg 27
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 72 CTGTACAAGTTGTGGTGATTGCGACCTGGGCGTGGGAAGACCATCATCATCAAGCGC 131
QY 28 PheThrAspAsnGluPheHisSerSerHisIleSerThrIleGlyValAspPheLysMet 47
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 132 TAGCTGCACCAAGAACTTCTCTCGCACTACCGGCCCAATCGCGTGGACTTCGCGCTC 191
QY 48 LysThrIleGluValAsp---GlyIleLysValArgIleGlnIleTyrAspThrAlaGly 66
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 192 AAGTGTCTCAGTGGGACCGGACGACTGTGGTGGCTCGAGCTCGGATATCGCAGGT 251
QY 67 GlnGluArgTyrGlnThrIleThrLysGlnTyrArgArgAlaGlnGlyIlePheLeu 86
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 252 CAAGAAAGATTGGAACATGACGAGGCTCTATTACCGAGAAGCTATGGGTGCATTATT 311
QY 87 ValTyrAspIleSerSerGluArgSerTyrGlnHisIleMetLysTrpValSerAspVal 106
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 312 GTCTTCGATGTCACGAGCGCCACATTTGAAGCTGTCGCAAGTGGAAATGATTG 371
QY 107 AspGlu-----TyrAlaProGluGly-----ValGlnLysIleLeuIleGlyAsnLys 122
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 372 GACTCCAAAGTAAGTCTCCCTAATGGCAACCGGTTTCAGTGGTTTGTGGCCACAAA 431
```

```
QY 123 AlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGly-----GlnGlnLeu 139
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 432 TGTGAC-----CAGGGGAAGGATGTCTCATGAACAATGGCCCTCAAGATGGACCATTC 485
QY 140 AlaLysGluTyrGly---MetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIle 158
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 486 TGAAGGAGCAGCGGTTCGTAGGATGCTTTGAAACATCAGCAAGAGAAATATAACATT 545
QY 159 LysGluSerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGlu 178
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 546 GATGAAGCCTCCAGATGCCTGGTCAACACATACTT----- 581
QY 179 GlyLeuArgMetArgAlaSerAsnGluLeuAlaLeuAlaGlu---LeuGluGluGluGlu 197
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 582 -----GCAATAGTGTGTGACCTAATGGATGCTATTGAGCCGCGACGTC 623
QY 198 GlyLysPro-----GluGlyProAla-AsnSerSe 207
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 624 GTGAAGCCCATCTCATCATCAACCAAGTTGCCAGTGTCTGTGGCTGTGCCAAATCCTAG 683
QY 207 rLys---ThrCysTrpCys 212
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 584 TAGGCACCTTGTGTGGTGT 702

RESULT 13
US-09-075-454-8
; Sequence 8, Application US/09075454
; Patent No. 6391580
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Tang, Y. Tom
; APPLICANT: Lal, Preeti
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Patterson, Chandra
; APPLICANT: Batra, Sajeev
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: RAS PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 6.1/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/075,454
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/766,551
; FILING DATE: DECEMBER 12, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0168-1 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1172 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 23, 2003, 11:10:12 ; Search time 158 Seconds
(without alignments)
1968.953 Million cell updates/sec

Title: us-09-817-198B-2

Perfect score: 1105

Sequence: 1 MAKQYDVLFRLLIGDVG.....LEEKGKPGPANSSKTCWC 212

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09817198/runat_18062003_144634_27974/app_query.fasta_1.391
-DB=Published Applications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62
-TRANS=human40.cdi -LIST=45 -LOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09817198_ECGN_1_1_57_erunat_18062003_144634_27974
-NCPU=6 -ICPU=3 -NO_MMAB -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1105	100.0	2021	9	US-09-764-868-88 Sequence 88, Appl
2	1105	100.0	3257	10	US-09-817-198A-1 Sequence 1, Appl1
3	832	75.3	566	9	US-09-764-868-507 Sequence 507, Appl
4	534.5	48.4	1274	10	US-09-925-302-91 Sequence 91, Appl

	5	530.5	48.0	1537	10	US-09-925-300-631	Sequence 631, App
	6	528.5	47.8	624	10	US-09-794-257-9	Sequence 9, Appl1
	7	528.5	47.8	1161	10	US-09-794-257-7	Sequence 7, Appl1
	8	528.5	47.8	2497	10	US-09-834-975-879	Sequence 879, App
	9	528.5	47.8	2497	10	US-09-834-975-885	Sequence 885, App
	10	528.5	47.8	2497	10	US-09-834-975-894	Sequence 894, App
	11	528.5	47.8	2497	10	US-09-834-975-896	Sequence 896, App
	12	519	47.0	651	9	US-09-938-842A-836	Sequence 836, App
	13	478.5	43.3	925	10	US-09-967-736-4	Sequence 4, Appl1
	14	478.5	43.3	939	9	US-10-102-806-48	Sequence 48, Appl1
	15	477.5	43.2	881	10	US-09-770-445-529	Sequence 529, App
	16	468	42.4	609	9	US-09-938-842A-832	Sequence 832, App
	17	456.5	41.3	1069	9	US-10-037-270-646	Sequence 646, App
	18	448	40.5	585	10	US-09-917-800A-1461	Sequence 1461, App
	19	438	39.6	446	9	US-09-918-995-22915	Sequence 22915, A
	20	429	38.8	875	12	US-10-051-986-10	Sequence 10, Appl
	21	429	38.8	1106	9	US-10-037-270-959	Sequence 959, App
	22	429	38.8	1116	10	US-09-794-257-13	Sequence 13, Appl
	23	429	38.8	2674	10	US-09-817-199A-1	Sequence 1, Appl1
	24	427	38.6	2623	9	US-09-764-868-71	Sequence 71, Appl1
	25	426	38.6	1316	9	US-09-764-868-493	Sequence 493, App
	26	424	38.4	833	9	US-10-037-270-426	Sequence 426, App
	27	420	38.0	429	9	US-09-918-995-5605	Sequence 5605, App
	28	419	37.9	639	10	US-09-350-874-66	Sequence 66, Appl
	29	416	37.6	576	10	US-09-794-257-15	Sequence 15, Appl
	30	415.5	37.6	894	10	US-09-770-445-487	Sequence 487, App
	31	412.5	37.3	538	10	US-09-924-035A-794	Sequence 794, App
	32	411.5	37.2	654	9	US-09-938-842A-2113	Sequence 2113, App
	33	410	37.1	1000	9	US-09-764-868-488	Sequence 488, App
	34	406	36.7	896	10	US-09-770-445-478	Sequence 478, App
	35	404.5	36.6	3936	10	US-09-919-172-49	Sequence 49, Appl
	36	399	36.1	1127	9	US-09-764-868-487	Sequence 487, App
	37	399	36.1	3124	10	US-09-925-302-340	Sequence 340, App
	38	395	35.7	844	10	US-09-770-445-646	Sequence 646, App
	39	391	35.4	1673	9	US-09-954-531-1359	Sequence 1359, App
	40	388.5	35.2	4083	10	US-09-817-182-1	Sequence 1, Appl1
	41	387.5	35.1	654	9	US-09-938-842A-78	Sequence 78, Appl
	42	387	35.0	771	10	US-09-728-445-652	Sequence 652, App
	43	384.5	34.8	1042	10	US-09-954-456-563	Sequence 563, App
	44	382.5	34.6	676	10	US-09-770-149-374	Sequence 374, App
	45	381	34.5	645	9	US-09-938-842A-774	Sequence 774, App

ALIGNMENTS

RESULT 1
US-09-764-868-88
; Sequence 88, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764, 868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 88
; LENGTH: 2021
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-868-88

Alignment Scores:
Pred. No.: 4.47e-132 Length: 2021
Score: 1105.00 Matches: 212
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-817-198B-2 (1-212) x US-09-764-868-88 (1-2021)

RESULT 5
US-03-925-300-631
; Sequence 631, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, P


```

Qy      81 AlaGInGlyLIlePheLeuValTyAspIleSerSerGluArgSerTyrGlnHisIleMet 100
Db      306 GCTCATGGCATTCGTGGTGTATGACGTCTACTGACCAAGGANTCCTACGCCACAGTGAAG 365

Qy      101 LysTrpValSerAspValAspGluTyTAlaProGluGlyValGlnLysIleLeuIleGly 120
Db      366 CAGTGGCTCAGGAGATTGACGGCTATGCCAGCGAGAAGCGTCAATAAGCTCCTGGTGGGC 425

Qy      121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlyGlnGlnLeuAla 140
Db      426 NACAAGAGCGACCTCACCACCAAGAAGGTGTGTGACCAACACCACAGCCAAAGGAGTTGCA 485

Qy      141 LysGluTyGlyMetAspPheTyrGlnThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
Db      486 GACTCTCTGGGCATCCCTTCTTGGAGACGAGCGCCCAAGATGCCACCAATGTCGAGCAG 545

Qy      161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu 180
Db      546 GCGTTC-----ATGACCATGGCTGCTGAAATCAAAAAGCGGATGGGCGCT 590

Qy      181 ArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGluGlyLysPro 200
Db      591 GGACGAGCCTCTGGGGGCGAGCGGCCCAATCTCAGATCGACACACCCCTGTAAAGCCG 650

Qy      201 GluGly 202
Db      651 GCTGGC 656

RESULT 14
US-10-102-806-48
; Sequence 48, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PAI03PIC1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
; LENGTH: 939
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (937)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-102-806-48

Alignment Scores:
Pred. No.: 6,3e-52 Length: 939
Score: 478.50 Matches: 91
Percent Similarity: 65.84% Conservative: 42
Best Local Similarity: 45.05% Mismatches: 64
Query Match: 43.30% Indels: 5
Db: 9 Gaps: 1

US-09-817-198B-2 (1-212) x US-10-102-806-48 (1-939)

Qy      1 MetalAlaLysGlnTyAspValLeuPheArgLeuLeuLeuLeuLeuLeuValGly 20
Db      70 ATGACCCCGAATAGACTACCTGTTTAACTGCTTTTGAATGGGACTCAGGCGTGGGC 120

Qy      21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSerThr 40
Db      130 AAGTCATGCTGCTCTCTCGGCTTTGCTGTATGACAGCTACACAGAGAGCTACATCAGACC 180

```



```

QY 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60
DB 190 ATCGGGTGGACTTCAAGATCCGACCATCGAGCTGGATGGCAAACTATCAAACTTCAG 249
QY 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
DB 250 ATCTGGGACACACGGCGCCAGGAGCGTCCGGACCATCATCTCCAGCTACTACCGGGG 309
QY 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
DB 310 GCTCATGGCATCATCGTGGTGTATGAGCTCACTGACGAGCAATCTTACGCCAAGTGAAG 369
QY 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
DB 370 CAGTGGCTGCGAGGAGATTGACCTTCTTGGAGACGAGCAACGCTCAATAGCTCTCGTGGGC 429
QY 121 AsnLysAlaAspGluGlnGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAla 140
DB 430 AACAGAGGACCTCACCCACCAAGAGGTGTGGACACACACACGCAAGGAGTTGCA 489
QY 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
DB 490 GACTCTCTGGGCATCCCTTCTTGGAGACGAGCGCCAAAGATGCCACCAATGTGCGAGCAG 549
QY 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu 180
DB 550 GCCTTC-----ATGACCATGGCTGCTGAATCAAAAGCGGATGGGCGCT 594
QY 181 ArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGluGluGlyLysPro 200
DB 595 GGAGCAGCCCTCTGGGGCGGAGCGGCCCATCTCAAGATCGACAGCACCCTCTGAAGCGG 654
QY 201 GluGly 202
DB 655 GCTGGC 660

```

RESULT 15

US-09-770-445-529
Sequence 529, Application US/09770445
Patent No. US2002023281A1

GENERAL INFORMATION:

```

; APPLICANT: Goralach, Jörn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krickler, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; TITLE OF INVENTION: thaliana
; FILE REFERENCE: 2023US (PARA-012PRV)
; CURRENT APPLICATION NUMBER: US/09/770,445
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,472
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 529
; LENGTH: 881
; TYPE: DNA

```

; ORGANISM: Arabidopsis thaliana
US-09-770-445-529

Alignment Scores:

```

Pred. No.: 7,72e-52 Length: 881
Score: 477.50 Matches: 96
Percent Similarity: 65.57% Conservatives: 43
Best Local Similarity: 45.28% Mismatches: 55
Query Match: 43.21% Indels: 18
DB: 10 Gaps: 4

```

US-09-817-198B-2 (1-212) x US-09-770-445-529 (1-881)

```

QY 1 MetalalysGlnTyrAspValLeuPheArgLeuLeuIleGlyAspSerGlyValGly 20
DB 60 ATGAATCCTCAGTACGACTATCTTTCAGCTCTCTTATCGGGATTCCTGGCGTAGGC 119
QY 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisIleSerThr 40
DB 120 AAGTCITGTCTCTTTTGAGATTCTCTGATGATTCTTATGAGAAAGTTTACATTAGCCT 179
QY 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60
DB 180 ATTGGAGTCGATTTAAATTAGGACTGTGGAACAAGATGCAAAACAATTAAGCTCAA 239
QY 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
DB 240 ATTTGGACACTGCTGCTCAAGAACGGTTCAGGACTATTACTAGCAGTTACTACCGTGG 299
QY 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
DB 300 GCACATGGAATTAATTATGTCTACGATGTCAAGATGCAAGAAAGCTTCAATATGTCAAG 359
QY 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
DB 360 CAATGGTTGAGTGAATGATCGTTATGCTAGTGACAAATGTCAACAACCTCCTTTGTGGA 419
QY 121 AsnLysAlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAla 140
DB 420 AACAAAGTCTGATCTTACTGAAAACAGAGCCATTCCTTATGAAACTGCCAAGGCTTTGCC 479
QY 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
DB 480 GATGAATCGGGATTCCTTTTATGGAGACTAGTCAAAAGATGCTACAAACGTTAGAACAG 539
QY 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu 180
DB 540 GCTTTC-----ATGGCAATGCTGCTCATCCATCAAGAG----- 572
QY 181 ArgMetArgAlaSer-AsnGluLeuAlaLeuAlaGluLeuGluGluGluGlyLysPr 200
DB 573 AGAATG---GCTAGCCCAACAGCTGGGAATAATGCAA-----GA 608
QY 200 oGluGlyProAlaAsnSerSerLysThrCysTrp 211
DB 609 CCACCGACCGTGCAGATCAGAGGACGCGCTGTGG 642

```

Search completed: June 23, 2003, 12:54:35

Job time : 161 secs

THIS PAGE BLANK (USPTO)